

Generalized Transform Analysis of Affine Processes And Asset Pricing Applications*

Hui Chen[†] Scott Joslin[‡]

June 24, 2009

Abstract

Nonlinearity is an important consideration in many problems of finance and economics, such as pricing securities, computing equilibrium, and conducting structural estimations. This paper significantly extends the transform analysis in Duffie, Pan, and Singleton (2000) by providing analytical treatment of a general class of nonlinear transforms for affine jump-diffusions. As example applications, we show how the generalized transform can be used to price defaultable bonds with state-dependent recoveries, and to solve for models of heterogeneous beliefs under much richer settings than considered in existing literature. These examples demonstrate the wide applicability of the generalized transform in economic modeling.

Keywords: affine process, generalized transform, tempered distributions, fourier transform, asset pricing, credit risk, heterogeneous beliefs

*We thank Ian Martin, Damir Filipovic, Ken Singleton, and the seminar participants at MIT Sloan, Boston University, and the 2009 Adam Smith Asset Pricing Conference for comments. Tran Ngoc-Khanh provided excellent research assistance. All the remaining errors are our own.

[†]MIT Sloan School of Management, huichen@mit.edu

[‡]MIT Sloan School of Management, sjoslin@mit.edu

1 Introduction

In this paper, we provide analytical treatment of a class of transforms for state variables that follow affine jump-diffusions (AJD). These transforms bring analytical and computational tractability to a large class of nonlinear moments, and can be useful in many applications in economics and finance. We apply the generalized transform to a variety of examples, including option pricing, term structure modeling, credit risk modeling, method of moment estimations, and computing the equilibrium of asset pricing models with multiple goods or heterogenous agents.

For a state variable X_t that follows an affine process, in the sense that the conditional characteristic function is affine¹, [Duffie, Pan, and Singleton \(2000\)](#), hereafter DPS, derive closed-form expression for the following transform:

$$E_t \left[\exp \left(- \int_t^T R(X_s, s) ds \right) e^{u \cdot X_T} (v_0 + v_1 \cdot X_T) 1_{\{\beta \cdot X_T < y\}} \right], \quad (1)$$

where $R(X)$ is an affine function of X , which can be interpreted as a stochastic “discount rate”, and $e^{u \cdot X_T} (v_0 + v_1 \cdot X_T) 1_{\{\beta \cdot X_T < y\}}$ is the terminal payoff function at time T .

We generalize the DPS result by deriving closed-form expression (up to an integral) for the following transform:

$$E_t \left[\exp \left(- \int_t^T R(X_s, s) ds \right) f(X_T) g(\beta \cdot X_T) \right], \quad (2)$$

where f can be a polynomial, a log-linear function, or the product of the two; g is a piecewise continuous function with at most polynomial growth (or more generally a tempered distribution) satisfying certain regularity conditions. When $f(X) = e^{u \cdot X} (v_0 + v_1 \cdot X)$ and $g(\beta \cdot X) = 1_{\{\beta \cdot X < y\}}$, we recover the transform of DPS in (1). The flexibility in choosing f and g in (2), as well as the state variable X_t , makes the generalized transform useful in dealing with generic nonlinearity problems in pricing (nonlinear stochastic discount factors or payoffs), estimation (nonlinear moments), and other areas of economic modeling.

The primary analytic tool that we use is the Fourier transform. In particular,

¹See [Duffie, Filipovic, and Schachermayer \(2003\)](#) for an elaboration on the characterization via the characteristic function.

we utilize knowledge of the conditional characteristic function of the state variable X_t (under certain forward measures) jointly with a Fourier decomposition of the nonlinearity in g . This combination brings tractability to our generalized transform by avoiding intermediate Fourier inversions.

The generalized transform method not only provides a convenient tool for computations, but also sheds further light on how we can take advantage of the special properties of affine processes in economic modeling. Consider the basic pricing equation for an asset with stochastic payoff Y_T at time T :

$$P_t = E_t [m_T Y_T],$$

where m_t is the stochastic discount factor. In the background, there is a model that determines (either exogenously or endogenously) the discount factor m and payoff Y as functions of the state variables X . In exchange for tractability, we often sacrifice the richness and realism of the model by adopting special utility functions or imposing strong restrictions on the processes of the state variables.

However, the new tools provided in this paper help relax some of these modeling constraints. By modeling X as an affine process and using the generalized transform, we can (i) price a wider range of assets with nonlinear payoffs, and (ii) significantly extend some existing models that produce nonlinear stochastic discount factors. In some cases these extensions are necessary to improve the quantitative performances of the models (e.g., by relaxing the assumptions of logarithmic preferences). Other times they provide a systematic and convenient way to introduce new elements to the existing models, such as time-varying growth rates, stochastic volatility, or jumps, as long as they can be summarized by the affine state variable X_t . We provide several example applications along these lines.

Option pricing. For many European options (including exotic options), the payoff function can be expressed as the product of exponential functions and an indicator function. When the underlying stock price follows an affine process, the option value can be computed in closed form through the transform.

Nonlinear Taylor Rule. We present a nonlinear Taylor Rule model that generalizes the model of fed funds target in Piazzesi (2005). We specify the policy function to maintain the requirement that the fed funds target rate, f_t , is non-negative, has increments in a multiple of 25 basis points, and depends on macro variables such as

GDP growth and inflation, but relax the restriction that the distribution of f_t have the exponential-affine Laplace transform. One application of such a model is to price fed funds futures.

Recovery Risk. The recovery rate of a corporate bond upon default has to be between 0 and 1. There are also empirical studies showing that recovery rates nonlinearly depend on macroeconomic, industry, and firm-specific variables. We introduce a class of state-dependent recovery models, relaxing the “recovery of market value” assumption standard in current literature. We derive closed form solutions for the pricing of defaultable zero-coupon bonds. The model can also be used to price other recovery-sensitive credit instruments such as credit default swaps or recovery locks². Our example of Cauchy recovery model demonstrates that ignoring the correlation between recovery rates and default intensities can lead to substantial deviations in credit spreads, especially for bonds with high or low credit quality.

Models of multiple goods or heterogeneous agents. Nonlinearity in the stochastic discount factors arise in many consumption-based asset pricing models. For example, in a Lucas model with multiple trees, the stochastic discount factor depends nonlinearly on the dividends from the trees. There are similar features in models of heterogeneous agents, including differences of beliefs, differences in risk aversion, and differences in the preferences for consumption goods (e.g., home bias in models of international trade). Using a model of two trees (Cochrane, Longstaff, and Santa-Clara (2008)) and a model of exchange rates (Pavlova and Rigobon (2007)) as examples, we show how to allow for more general preferences and introduce features such as mean reversion and conditional heteroscedasticity in dividend growth, as well as jumps with time-varying intensities in these models. We also solve a model of differences of beliefs about economic disasters and examine the equilibrium effects of the disagreements on asset prices.

Method of moments estimation. The need to compute unconditional and conditional moments of nonlinear functions also arises in the method of moments or GMM estimations. We provide a GMM estimator to a model with latent state variables.

Thanks to its tractability and flexibility, affine processes have been widely used in term structure models, reduced-form credit risk models, and option pricing. In

²This is a forward contract that requires no upfront or running payments, and allows purchase or sale of underlying bonds at a predetermined price if a credit event occurs.

particular, the transform analysis of general affine jump-diffusions in [Duffie, Pan, and Singleton \(2000\)](#) makes it easy to compute the moments arising from a wide range of pricing problems. Examples applications include [Singleton \(2001\)](#), [Pan \(2002\)](#), [Piazzesi \(2005\)](#), and [Joslin \(2009\)](#), among many others.

When the moment functions do not fit into the DPS transform, alternative methods to compute the moments include simulations or solving numerically the partial differential equations arising from the expectations via the Feynman-Kac methodology. Both methods can be time-consuming and lacking accuracy, especially in high dimensional cases, making them less suitable for estimations or doing comparative statics.

Yet another approach to compute the nonlinear moments of affine processes is to first recover the conditional density of the affine state variables through Fourier inversion of the conditional characteristic function, which in turn can be computed using the transform analysis of DPS, and is available in closed form in some special cases. Then, one can evaluate the nonlinear moments directly using the density. Through this method, [Duffie, Pan, and Singleton \(2000\)](#) obtain the extended transform in (1) for affine jump-diffusions, which they apply to option pricing. [Bakshi and Madan \(2000\)](#) connect the pricing of a class of derivative securities to the characteristic functions for a general family of Markov processes. Other papers that take this approach include [Heston \(1993\)](#), [Chen and Scott \(1995\)](#), [Bates \(1996\)](#), [Bakshi, Cao, and Chen \(1997\)](#), [Dumas, Kurshev, and Uppal \(2009\)](#), among others. A limitation of this approach is that, when computing general moments, it involves multiple numerical integrals and is subject to the curse of dimensionality.

Alternatively, we can take the Fourier transform of the moment function, effectively replacing the nonlinear moment function with an integral of exponential functions, for which the expectations can be computed analytically through the DPS transform. For example, [Carr and Madan \(1999\)](#) address the nonlinearity in a European option payoff by taking the Fourier transform of the payoff function with respect to the strike price. [Martin \(2008\)](#) takes the Fourier transform of a nonlinear pricing kernel that arises in the two tree model of [Cochrane, Longstaff, and Santa-Clara \(2008\)](#).³ In both cases, the conditional characteristic functions from the Fourier transform are known in closed form, which gives option prices or the value of consumption claims in almost closed form.

³In the N -tree case, $N > 2$, [Martin \(2008\)](#) also provides an $(N - 2)$ -dimensional integral to compute the associated $(N - 1)$ -dimensional transform.

To the best of our knowledge, this paper is the first to generalize this approach. We provide analytical treatment of generic nonlinear moments of affine processes by combining the Fourier method with the transform analysis of DPS. Our method shows the power of affine processes as a modeling workhorse not just in the “traditional fields” such as option pricing and term structure modeling, but many other areas of economics and finance.

2 Generalized Transforms

In this section, we outline our theoretical results. Full details are deferred to the Appendix. As in DPS, we begin by fixing a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and an information filtration $\{\mathcal{F}_t\}$, satisfying the usual conditions (see e.g., [Protter \(2003\)](#)), and suppose that X is a Markov process in some state space $D \subset \mathbb{R}^n$ satisfying the stochastic differential equation

$$dX_t = (K_0 + K_1 X_t)dt + \sqrt{H_0 + H_1 \cdot X_t}dW_t + dZ_t, \quad (3)$$

where W is an \mathcal{F}_t -standard n -dimensional Brownian motion and Z is a pure jump process with arrival intensity $\lambda_t = \lambda_0 + \lambda_1 \cdot X_t$ and fixed D -invariant distribution ν . Whenever needed, we also assume that there is an affine discount rate function $R(X_t) = \rho_0 + \rho_1 \cdot X_t$. For brevity, let Θ denote the parameters of the process $(K_0, K_1, H_0, H_1, \lambda_0, \lambda_1, \nu, \rho_0, \rho_1)$. Alternatively, we can define the process in terms of the infinitesimal generator or, as [Duffie, Filipovic, and Schachermayer \(2003\)](#) and [Singleton \(2001\)](#) stress, in terms of the conditional characteristic function.

2.1 Transform Analysis

In order to establish our main result, let us first review some basic concepts from distribution theory. A function $f : \mathbb{R}^N \rightarrow \mathbb{R}$ which is smooth and rapidly decreasing in the sense that for any multi-index α and any $N \in \mathbb{N}$, $\|f\|_{N,\alpha} \equiv \sup_x |\partial^\alpha f(x)|(1+|x|)^N < \infty$ is referred to as a Schwartz function. The collection of all Schwartz functions is denoted \mathcal{S} . \mathcal{S} is endowed with the topology generated by the family of semi-norms $\|f\|_{N,\alpha}$. The dual of \mathcal{S} , denoted \mathcal{S}^* and also called the set of tempered distributions, is the set of continuous linear functionals on \mathcal{S} . Any continuous function which has at most polynomial growth in the sense that $|g(x)| < |x|^p$ for some p and x large enough

is seen to be a tempered distribution through the map

$$g : \mathcal{S} \rightarrow \mathbb{R} \quad g : f \mapsto \int_{x \in \mathbb{R}^N} g(x)f(x)dx. \quad (4)$$

Many tempered distributions do not arrive from functions. An important example is the δ -function, $\delta : h \mapsto h(0)$. For our considerations, the key property is that the set of tempered distributions is suitable for Fourier analysis. For example, a function which is bounded may not have a Fourier transform in the sense of a function, but will possess a Fourier transform that is a tempered distribution. An example is the Heaviside function:

$$f(x) = 1_{\{x \leq 0\}} \Rightarrow \hat{f}(s) = \frac{1}{2}\delta(s) - \frac{1}{2\pi s}, \quad (5)$$

where integrating against $1/s$ is to be interpreted as the principle value of the integral. Considering distributions allows us to consider functions which are not integrable and thus in particular may not decay at infinity and may not even be bounded.

We now state our main result:

Theorem 1. *Suppose that $f(s) = \exp(s)$, $g \in \mathcal{S}^*$ and (Θ, α, β) satisfies Assumption 1 and Assumption 2 in [Appendix A](#). Then*

$$\begin{aligned} H(f, g, \alpha, \beta) &= E_0 \left[\exp \left(- \int_0^T R(X_u) du \right) f(\alpha \cdot X_T) g(\beta \cdot X_T) \right] \\ &= \frac{1}{2\pi} \langle \hat{g}, G(\alpha + \cdot \beta i) \rangle, \end{aligned} \quad (6)$$

where $\hat{g} \in \mathcal{S}^*$ and $G(\alpha + \cdot \beta i)$ denotes the function

$$s \mapsto G(\alpha + s\beta i) = E_0 \left[e^{-\int_0^T R(X_u) du} e^{(\alpha + is\beta) \cdot X_T} \right]. \quad (7)$$

The function G is the transform given in DPS. Recalling their result,

$$G(\alpha + is\beta) = e^{A(T; \alpha + is\beta, \Theta) + B(T; \alpha + is\beta, \Theta) \cdot X_0},$$

where A, B satisfy the ordinary differential equations (ODEs)

$$\dot{B} = K_1^\top B + \frac{1}{2}B^\top H_1 B - \rho_1 + \lambda_1(\phi(B) - 1) \quad B(0) = \alpha + is\beta, \quad (8)$$

$$\dot{A} = K_0^\top B + \frac{1}{2}B^\top H_0 B - \rho_0 + \lambda_0(\phi(B) - 1) \quad A(0) = 0, \quad (9)$$

where $\phi(c) = E_\nu[e^{c \cdot Z}]$, the moment-generating function of the jump distribution. Solving the ODE system (8–9) adds little complication to the transform. The solution is available in closed form in some cases, and can generally be quickly and accurately computed using standard numerical methods.

In the special case that \hat{g} defines a function, we can write the result as

$$H = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{g}(s) G(\alpha + is\beta) ds. \quad (10)$$

Fourier transforms of many functions are known in closed form.⁴ Additionally, standard rules allow for differentiation, integration, product, convolution and other operations to be conducted while maintaining closed form expressions. However, even if the function \hat{g} is not known in closed form, including those cases where g itself is given as an implicit function, it can be computed readily.

In some cases of interests, Assumption 2 may be violated. It could be that $\beta \cdot X_T$ has heavy tails so that, for example, $E[(\beta \cdot X_T)^4] = \infty$. Another example would be in a pure-jump process where the density may not be continuous. Depending on the case, our result can often be extended by limiting arguments or by considering different function spaces (such as Sobolev spaces for non-smooth densities).

There is some flexibility in the choice of α and g in (6). Notice that

$$e^{\alpha \cdot X_T} g(\beta \cdot X_T) = e^{(\alpha - c\beta) \cdot X_T} \tilde{g}(\beta \cdot X_T),$$

where $\tilde{g}(s) = e^{cs} g(s)$. This property can be useful in the case where g is not integrable but decreases rapidly as s approaches either $+\infty$ or $-\infty$ (e.g., the logit function). In this case, such a transformation of g makes it possible to apply (10).

⁴See, for example, [Folland \(1984\)](#).

2.2 Extensions of Generalized Transforms

The result above can be extended in a number of ways. First, we introduce a class of *pl-linear* (polynomial-log-linear) functions:

$$f(\alpha, v, X) = \sum_i p_i(v_i \cdot X) e^{\alpha_i \cdot X}, \quad (11)$$

where $\{p_i\}$ are arbitrary polynomials and $\{\alpha_i\}$ are complex vectors.⁵ We will refer to any function which cannot be expressed as the product of a polynomial and a log-linear function as *non-pl-linear*. The following proposition extends the function f in [Theorem 1](#) to any *pl-linear* functions.

Proposition 1. *Suppose that $v \in \mathbb{R}^N$, $n \in \mathbb{N}$, $g \in \mathcal{S}^*$ and (Θ, α, β) satisfies Assumption 1' and Assumption 2' in [Appendix B](#). Then*

$$\begin{aligned} H(f, g, \alpha, \beta, v, n) &= E_0 \left[\exp \left(- \int_0^T R(X_u) du \right) (v \cdot X_T)^n e^{\alpha \cdot X_T} g(\beta \cdot X_T) \right] \\ &= \frac{1}{2\pi} \langle \hat{g}, G(\alpha + \cdot \beta i; v, n) \rangle, \end{aligned} \quad (12)$$

where $\hat{g} \in \mathcal{S}^*$ and $G(\alpha + \cdot \beta i; v, n)$ denotes the function

$$s \mapsto G(\alpha + s\beta i; v, n) = E_0 \left[e^{-\int_0^T R(X_u) du} (v \cdot X_T)^n e^{(\alpha + is\beta) \cdot X_T} \right]. \quad (13)$$

The function G is computed by solving the associated ODE in [Appendix B](#).

The assumption that the function g in the generalized transform be a tempered distribution might appear restrictive at first sight, since g cannot have exponential growth (see our earlier discussions of Schwartz functions). However, as [Proposition 1](#) demonstrates, by specifying f and g appropriately, we can let f “absorb” any exponential (and polynomial) growth in a moment function, rendering g admissible to the transform. We will demonstrate this feature in several examples.

The transform in [Theorem 1](#) assumes that g can only depend on X through the linear combination $\beta \cdot X$. Thus, the marginal impact of X_i on g will be proportional to β_i , which might be too restrictive in some cases. The following proposition relaxes this restriction by considering $g(\beta_1 \cdot X, \dots, \beta_M \cdot X)$ for $M \in \mathbb{N}$.

⁵Allowing complex eigenvalues allows one to have oscillatory *sine* and *cosine* terms.

Proposition 2. Suppose that $f(s) = \exp(s)$, $g \in \mathcal{S}_M^*$ (an M -dimensional tempered distribution), $\alpha \in \mathbb{R}^N$, $\mathbf{b} \in \mathbb{R}^{M \times N}$ and $(\Theta, \alpha, \mathbf{b})$ satisfies Assumption 1 and Assumption 2 in [Appendix A](#). Then

$$\begin{aligned} H(f, g, \alpha, \mathbf{b}) &= E_0 \left[\exp \left(- \int_0^T R(X_s) ds \right) f(\alpha \cdot X_T) g(\mathbf{b} X_T) \right] \\ &= \left(\frac{1}{2\pi} \right)^M \langle \hat{g}, G_M(\alpha + \cdot \mathbf{b} i) \rangle, \end{aligned} \quad (14)$$

where $\hat{g} \in \mathcal{S}^*$ and $G_M(\alpha + \cdot \mathbf{b} i)$ denotes the function

$$G_M : \mathbb{C}^M \rightarrow \mathbb{C}, \quad s \mapsto G_M(\alpha + s^\top \mathbf{b} i) = E_0 \left[e^{-\int_0^T R(X_u) du} e^{(\alpha + i s^\top \mathbf{b}) \cdot X_T} \right]. \quad (15)$$

Finally, it is immediate to extend the transform in [Proposition 2](#) by replacing $f(\alpha \cdot X)$ with a pl -linear function as in [Proposition 1](#).

3 Applications

In this section, we outline some example application of the generalized transform in economic modeling. We emphasize how to formulate these problems to take advantage of the flexibility of affine processes and the tractability of the transform.

3.1 Option Pricing

As shown in DPS, for pricing European options, we want to evaluate the transform:

$$E_t^Q \left[e^{-\int_t^T r_s ds + \alpha \cdot X_T} g_y(\beta \cdot X_T) \right], \quad (16)$$

where $g_y(x) = 1_{\{x \leq y\}}$ is nonlinear and non-integrable. For example, for an European put option with strike K , X_t will be the log stock price, $y = \log K$, and the option price is

$$P_t = E_t^Q \left[e^{-\int_t^T r_s ds + y} g_y(X_T) \right] - E_t^Q \left[e^{-\int_t^T r_s ds + X_T} g_y(X_T) \right]. \quad (17)$$

However, the Fourier transform of g_y is defined as a distribution:

$$\hat{g}_y(s) = \frac{1}{2} \delta(s) + \frac{e^{i\pi y s}}{2\pi i s}, \quad (18)$$

where δ , the dirac- δ function, is the distribution defined by the relation

$$\int \delta(x)h(x)dx = h(0). \quad (19)$$

After applying symmetry, this replicates the formula given in DPS obtained by Levy-inversion.⁶

Since the generalized transform only requires the state variable X_t to be affine, we can introduce features such as stochastic interest rate, stochastic volatility, and jumps into the model of stock prices. Our method also applies to exotic options whose payoffs are more complicated functions of the terminal value of the stock (and other state variables).

3.2 Nonlinear Taylor Rule

A common approach to modeling the monetary policy rule of the Fed is in the form of a simple Taylor rule (Taylor (1993)),

$$f_t = \beta_0 + \beta_1\pi_t + \beta_2g_t + \epsilon_t, \quad (20)$$

where f_t denotes the fed funds target rate at time t , π_t denotes inflation, g_t represents a measure of the output gap, and ϵ_t is the monetary policy shock. This gives a simple representation of the Fed's goal of price stability and sustainable economic growth.

Piazzesi (2005) uses pure jump processes with deterministic jump times to model the fed funds target f_t . Let X_t be a vector summarizing economic conditions, which may contain inflation and growth measures, as in a Taylor rule, and possibly other macroeconomic variables. In this specification, the distribution of the jump size does not depend on the state variable. This implies that the moment-generating function, conditional on pre-meeting information, maintains an affine form

$$E_{t_0-}[e^{af_t}] = e^{A_a+B_a \cdot X_{t_0-}}. \quad (21)$$

⁶In DPS, they arrive at this equation by effectively computing the forward density by Fourier transform (a 1-dimensional integral) and then integrating over the payoff region (now a 2-dimensional integral). In this case, Fubini and limiting arguments allow this 2-dimensional integral to be reduced to a 1-dimensional integral as in the standard Lévy inversion formula (without a forward measure).

This assumption implies, among other things, that the expected policy rule is linear in the state. In other words, in expectation a linear Taylor rule holds in the case that the state variables are the Taylor rule inputs, but there may be non-normal policy shock deviations from the linear rule.

The generalized transform technique allows us to consider more flexible nonlinear policy rules. Define a policy rule $G(\beta \cdot X_T)$ so that at a meeting at date T , the fed sets a target of $G(\beta \cdot X_T)$. The function G can be chosen such that it generates movements in f_t in a multiple of 25 basis points. Suppose the short rate is r_t , and the spread between the short rate and the target is $s_t = r_t - f_t$. Consider a federal funds futures contract which, for simplicity, we assume pays off f_T at some future FOMC meeting time T . The price of such a contract is given by:

$$P_t = E_t^Q \left[\exp \left(- \int_t^T r_s ds \right) f_T \right] = E_t^Q \left[\exp \left(- \int_t^T r_s ds \right) G(\beta \cdot X_T) \right]. \quad (22)$$

This expectation is easily mapped to the generalized transform (6), with

$$\begin{aligned} f(\alpha \cdot X) &= 1, \\ g(\beta \cdot X) &= G(\beta \cdot X_T). \end{aligned}$$

3.3 GMM Estimation

In the method of moments estimations, we often need to compute unconditional or conditional nonlinear moments. Consider an econometric model given by

$$u_t = y_t - h(X_t, w_t; \theta), \quad (23)$$

where θ is a vector of unknown parameters, X_t is a vector of latent state variables that follows a stationary affine process, y_t and w_t are variables observed at time t . For example, $h(X_t, w_t; \theta)$ can be the price of a security at time t given latent state variable X_t , observable state variable w_t , and parameters θ , while y_t can be the observed price of this security in the market. The moment condition is

$$E[u_t z_t] = 0, \quad (24)$$

for any valid instrument z_t .

A generalized method of moments estimator $\widehat{\theta}_{GMM}$ minimizes the objective function

$$Q(\theta; Y_T) = \left[\frac{1}{T} \sum_{t=1}^T z_t [y_t - E[h(X_t, w_t; \theta)]] \right]' W_T \left[\frac{1}{T} \sum_{t=1}^T z_t [y_t - E[h(X_t, w_t; \theta)]] \right], \quad (25)$$

where W_T is the weighting matrix. Assuming that h can be decomposed into the product of f and g satisfying the regularity conditions in [Theorem 1](#), then we can compute the expectation using the generalized transform.

3.4 Two Non-IID trees

[Cochrane, Longstaff, and Santa-Clara \(2008\)](#) show that in a Lucas economy ([Lucas \(1978\)](#)) with two trees, the equilibrium conditions imply rich dynamics for stock returns and volatility in the time series and cross section. They provide closed-form solutions with the assumption of log utility and *i.i.d.* trees. [Martin \(2008\)](#) extends the analysis to multiple trees, Poisson jumps in dividends, and power utility, but also assume that the dividend growth of each tree is *i.i.d.* As we outline in this example, the generalized transform can help substantially enrich the dynamics of the dividend processes.

There is an infinitely-lived representative investor with CRRA utility:

$$U_t = E_t \left[\int_0^\infty e^{-\rho u} \frac{C_{t+u}^{1-\gamma} - 1}{1-\gamma} du \right]. \quad (26)$$

There are two assets in the economy, both in unit supply, with dividend streams $D_{1,t}dt$ and $D_{2,t}dt$. We deviate from [Cochrane, Longstaff, and Santa-Clara \(2008\)](#) and [Martin \(2008\)](#) by considering *non-i.i.d* dividend processes. Specifically, we now suppose that there are time-varying conditional moments in the dividend processes in the form of variation in expected growth rates, volatilities, or disaster probabilities.

The log dividends $d_{1,t} = \log D_{1,t}$ and $d_{2,t} = \log D_{2,t}$ are driven by the following processes:

$$dd_{i,t} = g_{i,t}dt + \sigma_{d,i}dW_{i,t}^d + dZ_{i,t} \quad (27a)$$

$$dg_{i,t} = \kappa_{g,i}(\bar{g}_i - g_{i,t})dt + \sigma_{g,i}dW_{i,t}^g \quad (27b)$$

$$d\lambda_{i,t} = \kappa_{\lambda,i}(\bar{\lambda}_i - \lambda_{i,t})dt + \sigma_{\lambda,i}\sqrt{\lambda_{i,t}}dW_{i,t}^\lambda \quad (27c)$$

The term Z_i is a pure jump process, with arrival intensity $\lambda_{i,t}$ which follows a square-root process with long run mean $\bar{\lambda}_i$, and time-invariant distribution ν_i ; $g_{i,t}$ is the expected growth rate of log dividend $d_{i,t}$, which follows an Ornstein-Uhlenbeck process with long run mean \bar{g}_i . For simplicity, we assume that all the Brownian motions $W_{i,t}^d$, $W_{i,t}^g$, and $W_{i,t}^\lambda$ are uncorrelated with each other.

This model nests the two-tree model of [Cochrane, Longstaff, and Santa-Clara \(2008\)](#), which corresponds to the special case when $\gamma = 1$, $g_{i,t} \equiv \bar{g}_i$, and $\lambda_{i,t} \equiv 0$. When $\gamma > 1$ and the growth rates and disaster probabilities are constant ($g_{i,t} \equiv \bar{g}_i$ and $\lambda_{i,t} \equiv \lambda$), we recover the two-tree model of [Martin \(2008\)](#).

The equilibrium condition is that aggregate consumption $C_t = D_{1,t} + D_{2,t}$, which implies that the unique pricing kernel in this economy should be $M_t = e^{-\rho t} (D_{1,t} + D_{2,t})^{-\gamma}$. Under the standard regularity conditions, the price of stock i ($i = 1, 2$), $P_{i,t}$, is given by:

$$\begin{aligned} P_{i,t} &= E_t \left[\int_0^\infty \frac{M_{t+u}}{M_t} D_{i,t+u} du \right] \\ &= (D_{1,t} + D_{2,t})^\gamma \int_0^\infty e^{-\rho u} E_t \left[\frac{D_{i,t+u}}{(D_{1,t+u} + D_{2,t+u})^\gamma} \right] du. \end{aligned} \quad (28)$$

The main challenge of computing the stock prices comes from the stochastic discount factor, which is non-*pl-linear* in the log dividends. To map the expectation in (28) into the generalized transform, we first define the state variable $X_t = [d_{1t} \ g_{1t} \ \lambda_{1t} \ d_{2t} \ g_{2t} \ \lambda_{2t}]'$. It is easy to verify that X is affine:

$$dX_t = (K_0 + K_1 X_t) dt + \sqrt{H_0 + H_1 \cdot X_t} dW_t + dZ_t, \quad (29)$$

where K_0 , K_1 , H_0 , H_1 are implied by (27a)-(27c). Next, the expectation inside the integral of (28) can be rewritten as

$$E_t \left[\frac{D_{1,s}}{(D_{1,s} + D_{2,s})^\gamma} \right] = E_t \left[\frac{e^{(1-\gamma/2)d_{1,s} - \gamma/2 d_{2,s}}}{\left(2 \cosh \frac{d_{1,s} - d_{2,s}}{2} \right)^\gamma} \right] = E_t [f(\alpha \cdot X_s) g(\beta \cdot X_s)], \quad (30)$$

where

$$\begin{aligned} f(x) &= e^x, \\ g(x) &= \frac{1}{(2 \cosh(x))^\gamma}, \end{aligned}$$

and

$$\alpha = \left[\left(1 - \frac{\gamma}{2}\right) \ 0 \ -\frac{\gamma}{2} \ 0 \right]^T, \quad \beta = \left[\frac{1}{2} \ 0 \ -\frac{1}{2} \ 0 \right]^T.$$

Since X is affine and $g \in \mathcal{S}^*$, [Theorem 1](#) readily applies to [\(30\)](#). When the increments of X are *i.i.d.*, the conditional characteristic function for X is known explicitly, which [Martin \(2008\)](#) uses to compute [\(30\)](#) following a Fourier transform for g .

To compute the expected excess returns and volatilities of the stocks, we can consider each stock as a portfolio of zero-coupon equities, each of which with one dividend payment $D_{i,t}dt$. The values of these zero-coupon equities follow immediately from equation [\(30\)](#). The risk premium of the stock is then the value-weighted average of the risk premium for these zero-coupon equities, which can be easily computed by applying $\hat{\text{Ito}}$'s Lemma to the zero-coupon equity prices and the pricing kernel M_t .

This model can be further extended in several dimensions. First, within the affine framework, we can introduce additional state variables to X , such as stochastic volatility, or impose cointegration between the two dividend processes. Due to the time-separable utility function, these additional state variables do not directly enter into the pricing equation [\(28\)](#). Hence, adding these new state variables do not increase the dimension of the Fourier transform. Second, we can generalize the utility function, e.g., by making aggregate consumption C_t a CES aggregator of $D_{1,t}$ and $D_{2,t}$, as in [Piazzesi, Schneider, and Tuzel \(2007\)](#), where the two trees are interpreted as nonhousing consumption and housing services. It is also convenient to add preference shocks that are *pl-linear* in the state variables. Third, the model can allow for multiple trees using the multi-dimensional version of the generalized transform in [Proposition 2](#).

3.5 Models of Heterogeneous Agents

The next example illustrates how to use the generalized transform to solve models of heterogeneous agents with complete markets. The stochastic discount factors in the general form of these models will be implicit functions of the state variables. However, even in these general cases, the transform can still be applied by carefully choosing

the nonlinear function g .

We consider a generalized version of the model of international finance in [Pavlova and Rigobon \(2007\)](#).⁷ As in the two-tree model, there are two assets in the economy, with dividend streams $D_{1,t}dt$ and $D_{2,t}dt$, which can be defined the same way as (27a)–(27c), or as part of a more general affine process. Departing from the two-tree model, we introduce two agents A and B , who derive utility from the two goods:

$$V_i = E \left[\int_0^T e^{-\rho t} u_i(\bar{C}_{i,t}) dt \right], \quad i = A, B \quad (31)$$

where $u_i(c) = \frac{c^{1-\gamma_i}}{1-\gamma_i}$, and

$$\bar{C}_i = [\delta_i C_{i,1}^{\alpha_i} + (1 - \delta_i) C_{i,2}^{\alpha_i}]^{\frac{1}{\alpha_i}}, \quad (32)$$

where $\alpha_i = (\varepsilon_i - 1) / \varepsilon_i$, and ε_i represents agent i 's intratemporal elasticity of substitution between the two goods. In the international setting, as in [Pavlova and Rigobon \(2007\)](#), $\delta_i \in (0, 1)$ can be interpreted as the home bias in each country's (agent's) consumption basket. For simplicity, let us consider the case where $\alpha_A = \alpha_B$, $\gamma_A = \gamma_B$, $\delta_A = 1 - \delta_B$ (that is, the same degree of home bias) and simply drop the corresponding subscripts.

This model generalizes [Pavlova and Rigobon \(2007\)](#) in two aspects. As in the previous example, we can model dividends as part of a general affine jump-diffusion, as opposed to having *i.i.d.* growth. In addition, [Pavlova and Rigobon \(2007\)](#) assume that agents have log utility ($\gamma_A = \gamma_B = 1$) and unit intratemporal elasticity of substitution ($\alpha_A = \alpha_B = 0$), which have important implications on asset prices and risk premia in their model. We relax these restrictions, and more generally can allow for heterogeneity in preferences across agents.

The market clearing conditions are

$$C_{A,1} + C_{B,1} = D_1, \quad (33a)$$

$$C_{A,2} + C_{B,2} = D_2. \quad (33b)$$

⁷Their model also have demand shocks that are martingales. For simplicity, we drop the demand shocks in this example, which do not affect the tractability of the model.

Assuming that market is complete, we set up the social planner's problem,

$$\max_{\{C_{A,1}, C_{A,2}, C_{B,1}, C_{B,2}\}} E \left[\int_0^T e^{-\rho t} \{u_A(\bar{C}_{A,t}) + \lambda u_B(\bar{C}_{B,t})\} dt \right],$$

subject to the market clearing conditions. In [Appendix C](#), we show that the stochastic discount factor of Agent A can be written as

$$\xi_A(t) = e^{-\rho t} e^{\frac{-1-\gamma+\alpha}{2}d_{1,t} - \frac{1-\gamma-\alpha}{2}d_{2,t}} g(d_{2,t} - d_{1,t}), \quad (34)$$

where $d_{i,t} = \log(D_{i,t})$, and g is an implicit function given in the appendix. The coefficients for $d_{1,t}$ and $d_{2,t}$ in the exponential term are chosen to ensure that g is a smooth and bounded function, so that [Theorem 1](#) can be applied.

Although notationally more cumbersome, the fact that g is given implicitly presents little added complication. The function g depends only on the preference parameters (α, δ, γ) and not the parameters governing the dynamics of the supply of the goods given in [\(27a–27b\)](#), which also simplifies any calibration exercise. Additionally, by the implicit function theorem, we can compute derivatives of g in terms of implicit functions.

4 Two Concrete Examples

In this section, we provide in-depth analysis of two examples, one on pricing defaultable bonds with state-dependent recovery, the other on difference-of-opinion models, to illustrate how the generalized transform is applied and how it helps significantly enrich the models.

4.1 Recovery Risk

The value of a credit-risky security (e.g., defaultable bonds or credit default swaps) depends on the discount rate, default probability, and the recovery value of the security in the event of default. Recovery risk refers to the uncertainty about the recovery rate. Due to the great amount of difficulty in forecasting the recovery rate far ahead of the default event, academics and practitioners have often treated recovery risk as a secondary consideration. We introduce a new class of stochastic recovery models and study the impact of state-dependent recovery risks on pricing.

We fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and two filtrations $\{\mathcal{F}_t : t \geq 0\}$, $\{\mathcal{G}_t : t \geq 0\}$. Following the reduced-form models, the default time is assumed to be a totally inaccessible \mathcal{G} -stopping time $\tau : \Omega \rightarrow (0, +\infty]$. For simplicity, we assume that under the risk neutral measure \mathbb{Q} , τ is doubly-stochastic driven by the filtration $\{\mathcal{F}_t : t \geq 0\}$, with intensity λ_t .⁸ The instantaneous riskfree rate is r_t .

Consider a T year defaultable zero-coupon bond with face value of 1, and the recovery value at default φ is a bounded predictable process that is adapted to the filtration $\{\mathcal{F}_t : t \geq 0\}$. The price of the bond is:

$$\begin{aligned} V_0 &= E^{\mathbb{Q}} \left[e^{-\int_0^\tau R_u du} \mathbf{1}_{\{\tau \leq T\}} \varphi_\tau \right] + E^{\mathbb{Q}} \left[e^{-\int_0^T R_u du} \mathbf{1}_{\{\tau > T\}} \right] \\ &= E^{\mathbb{Q}} \left[\int_0^T e^{-\int_0^t (R_u + \lambda_u) du} \lambda_t \varphi_t dt + e^{-\int_0^T (R_u + \lambda_u) du} \right]. \end{aligned} \quad (35)$$

The second equality follows from the doubly-stochastic assumption and certain regularity conditions (see [Duffie \(2005\)](#) for details).

[Duffie and Singleton \(1999\)](#) discuss three types of recovery models:

1. “recovery of treasury” (RT):

$$\varphi_t = (1 - L_t) P_t, \quad (36)$$

where P_t is the price at time t of an otherwise equivalent default-free bond; L_t is a value between 0 and 1.

2. “recovery of face value” (RFV):

$$\varphi_t = (1 - L_t) F, \quad (37)$$

where F is the face value of the bond.

3. “recovery of market value” (RMV):

$$\varphi_t = (1 - L_t) V_{t-}, \quad (38)$$

where V_{t-} is the market value of the security immediately before default.

⁸See [Duffie \(2005\)](#) for a survey on the reduced form approach for modeling credit risk and the doubly-stochastic property.

Duffie and Singleton (1999) show that under the RMV specification and a suitable no-jump condition,⁹ one can price defaultable claims with the “default-adjusted discount rate”, $r_t + \lambda_t L_t$. Moreover, if one directly specifies the mean-loss rate $\lambda_t L_t$ as affine, then the standard results for affine term structure models can be used to price defaultable bonds. In contrast, the RT and RFV models are generally less tractable.

While analytically appealing, the RMV assumption has some limitations. First, since the default intensity λ_t and the recovery rate L_t enter the default-adjusted discount rate symmetrically, we cannot separately identify the effect of default intensity and recovery using information on prices alone. Second, when pricing bonds of different seniorities from the same issuer, it is more natural to separately model default intensity (which is the same across different bonds) and recovery rates (which depends on seniority). Third, data on recovery rates are usually quoted as fraction of face value instead of market value. For example, Moody’s database of corporate defaults estimates defaulted debt recovery rates using the ratio of market bid prices (30 days after the date of default) to par value.

Bakshi, Madan, and Zhang (2006) study a class of RFV and RT models for which φ is exponential affine in the default intensity as well as the class of completely monotone functions.¹⁰ They solve for bond prices using the DPS transform. We show that a much wider range of RFV and RT models becomes tractable using the generalized transform analysis. In particular, we examine how the dependence of recovery rates on macro and firm-specific variables affect pricing.

4.1.1 Model Setup

We directly specify the dynamics of state variables $X_t = [\lambda_t \ Y_t]'$ under the risk neutral probability measure \mathbb{Q} :

$$d\lambda_t = \kappa_\lambda(\theta_\lambda - \lambda_t)dt + \sigma_\lambda \sqrt{\lambda_t} dW_t^\lambda, \quad (39)$$

$$dY_t = \kappa_Y(\theta_Y - Y_t)dt + \sigma_Y \sqrt{\lambda_t} dW_t^Y, \quad (40)$$

⁹The no-jump condition is satisfied here by assuming φ is adapted to $\{\mathcal{F}_t\}$. See also Duffie, Schroder, and Skiadas (1996) and Collin-Dufresne, Goldstein, and Hugonnier (2004) for discussions on the no-jump condition.

¹⁰The completely monotone class requires that each order derivative is either always positive or negative. This precludes, for example, an inflection point as in the Cauchy example we consider. Importantly, the Fourier transform provides for direct inversion formulas not available in the completely monotonic case.

where W_t^λ and W_t^Y are uncorrelated Brownian motions; λ_t is the default intensity of a firm, and the short term interest rate, r_t , is given by

$$r_t = Y_t - \delta\lambda_t. \quad (41)$$

This simple setup (with $\delta > 0$) captures the negative correlation between r_t and λ_t as argued by [Longstaff and Schwartz \(1995\)](#) and documented by [Duffee \(1998\)](#).

The recovery value φ of a bond issued by the firm can depend on the default intensity, the short rate, and other macro and firm-specific variables. For example, [Altman, Brady, Resti, and Sironi \(2005\)](#) document significant negative correlation between aggregate default rates and recovery rates. [Chen \(2008\)](#) provides evidence that macro variables such as GDP growth and riskfree rate are correlated with the aggregate recovery rates and default rates. Moreover, using a structural model, [Chen \(2008\)](#) shows that such comovements in default probabilities, recovery rates, and risk premia can have large effects on the pricing of defaultable bonds.

Another property for the recovery function φ is that, in principle, it should only take values from $[0, 1]$. One specification for φ that satisfies this requirement and is compatible with the DPS formulation is

$$\varphi(X) = e^{\beta \cdot X} 1_{\{\beta \cdot X < 0\}} + 1_{\{\beta \cdot X > 0\}}.$$

[Bakshi, Madan, and Zhang \(2006\)](#) study such a setting. More generally, the cumulative distribution function of any distribution will take values in $[0, 1]$. Below are some commonly used examples.

- Logit Model:

$$\varphi(X) = \frac{1}{1 + e^{-\beta_0 - \beta_1 \cdot X}}.$$

- Probit Model:

$$\varphi(X) = \int_{-\infty}^{\beta_0 + \beta_1 \cdot X} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds.$$

- Cauchy Model:

$$\varphi(X) = \int_{-\infty}^{\beta_0 + \beta_1 \cdot X} \frac{\gamma}{\pi((s^2 - s_0^2) + \gamma^2)} ds.$$

Modeling φ with CDFs has the additional benefit that they have nice Fourier transform properties. For example, the integrands of the Probit and Cauchy model have closed-

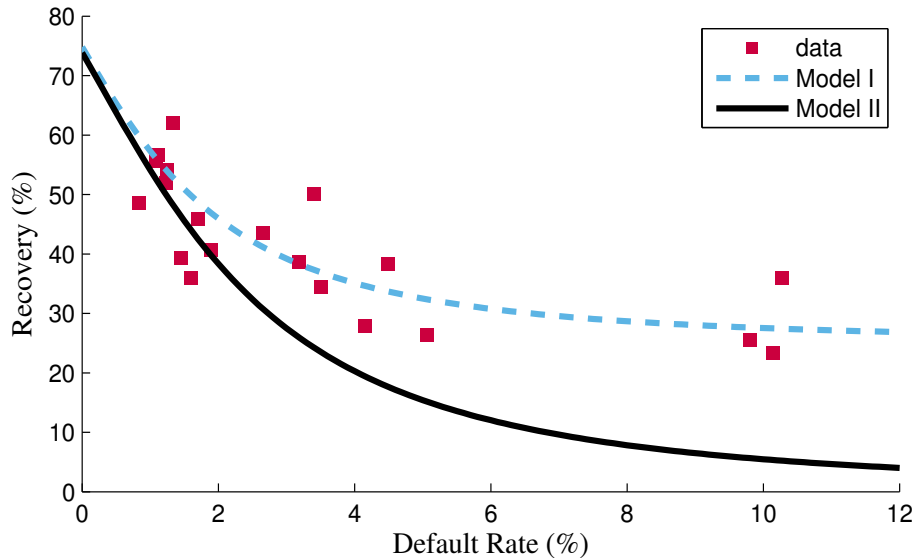


Figure 1: **A Cauchy Model of Aggregate Recovery Rates.** This figure plots the aggregate recovery rates and default rates from Altman et al. (2005). The dash line is a Cauchy recovery model fitted to the historical data. The solid line is a Cauchy recovery model with recovery risk premium.

form Fourier transform. Since Fourier transform has the property that $\hat{f}'(t) = t\hat{f}(t)$, it is very easy to obtain the Fourier transform of φ in those cases.

For simplicity, we assume that φ only depends on the default intensity, and we adopt a variation of the Cauchy model:

$$\varphi(\lambda) = \frac{a}{1 + b(\lambda - \lambda_0)^2} + c. \quad (42)$$

The constant term c sets a lower bound for φ . Its Fourier transform (excluding the constant c) is

$$\hat{\varphi}(t) = \frac{a\pi}{\sqrt{b}} e^{\lambda_0 it - \frac{1}{\sqrt{b}}|t|}. \quad (43)$$

We consider two calibrations of $\varphi(\lambda)$ in (42). First, using data from Altman et al. (2005), we calibrate $a = 0.68$, $b = 2000$, $c = 0.25$, and $\lambda_0 = -0.014$. The fitted function is “Model I” in Figure 1. The fitted curve is downward sloping and convex. The recovery rate is close to 70% when the probability of default is very low. When annual default probability rises to 10%, the recovery rate drops to 30%. The parametrization of Model I is likely too conservative: it treats the recovery rates in the data the

same as the risk-neutral recovery rates, assuming no recovery risk premium. In the second calibration, we assume $a = 0.9$, $b = 1200$, $c = 0$, and $\lambda_0 = -0.014$. The fitted function is “Model II” in [Figure 1](#), which has very similar recovery rates to Model I when default intensity is low, but has a sharper decline in recovery rates than Model I when default intensity rises. The widening gap between the two models implies that the recovery risk premium in Model II is increasing with the aggregate default probabilities.

Several features of the recovery curve will matter for bond pricing: how fast (slope) and how far (right tail) the recovery rate drops with default rate, and how much curvature the recovery function has. We will investigate how each of these features affects pricing.

The key step in computing the value of the defaultable zero-coupon bond is to compute the expectation

$$E_0^Q \left[\exp \left(- \int_0^t (r_u + \lambda_u) du \right) \lambda_t \varphi(\lambda_t) \right],$$

which is mapped into the generalized transform of [Theorem 1](#) by choosing

$$\begin{aligned} f(\alpha \cdot X) &= \iota_1 \cdot X, \\ g(\beta \cdot X) &= \frac{a}{1 + b(\iota_1 \cdot X - \lambda_0)^2}, \end{aligned}$$

where $\iota_1 = [1 \ 0]'$.

It is straightforward to use the recovery model $\varphi(X)$ to price other credit products, such as credit default swaps or recovery locks. In addition, our model can be generalized to allow for violations of the no-jump conditions.¹¹ Thus, it can be used in models with flight-to-quality, default contagion, systematic jump risk, or other features that violate the no-jump condition.

4.1.2 Results

We now use the processes of default intensity λ_t and riskfree rate r_t ([39–41](#)) and the recovery model ([42](#)) to price a 5-year defaultable zero-coupon bond of a representative firm, which has the same default intensity as the aggregate intensity λ_t . We calibrate

¹¹We can either explicitly make the correction for jumps as in [Duffie, Schroder, and Skiadas \(1996\)](#), or use the change-of-measure method in [Collin-Dufresne, Goldstein, and Hugonnier \(2004\)](#).

Table 1: CALIBRATION OF THE RISK-NEUTRAL DYNAMICS OF λ AND Y

κ_λ	θ_λ	σ_λ	κ_Y	θ_Y	σ_Y	δ
-0.035	-0.08	0.07	0.02	0.10	0.06	0.1

the process of λ_t and Y_t under the risk-neutral measure following [Duffee \(1999\)](#). The parameter values are reported in [Table 1](#). Notice that $\kappa_\lambda < 0$, which is consistent with Duffee’s finding that the default intensity of a typical firm is nonstationary under the risk-neutral measure.

The results are reported in [Figure 2](#). The top panels investigate Stochastic Recovery Model I, where the recovery function is fitted to the historical recovery rates (no recovery risk premium); the bottom panels investigate Stochastic Recovery Model II, which assumes that the recovery risk premium increases with aggregate default intensity (see [Figure 1](#)). A popular assumption for default recovery in both academic analysis and industry practice is to assume 75% constant loss rate (see e.g., [Pan and Singleton \(2008\)](#)). This value is higher than the historical mean loss rate, which is a parsimonious way to capture the recovery risk premium. Thus, for comparison, we also report the results for two alternative models: (1) recovery of market value (RMV), with constant loss rate $L = 0.75$; (2) recovery of face value (RFV), with constant loss rate $L = 0.75$.

Panel A shows that, without the recovery risk premium, the credit spreads generated by the stochastic recovery model are lower than the RMV and RFV model with constant recovery rate. Panel B reports the pricing errors of the RMV and RFV model relative to the stochastic recovery model (assuming the latter is the true model), computed as

$$\frac{\text{stochastic recovery yield} - \text{constant recovery yield}}{\text{stochastic recovery yield} - \text{default-free yield}}.$$

As expected, if there is no recovery risk premium (as in Model I), the main concern of the constant recovery assumption is underpricing, i.e. they generate credit spreads that are too high. The pricing errors are large. For example, the pricing errors of the RMV model are at least 20%, and can be over 40% when default intensity is low ($< 2\%$).

The results are quite different for Stochastic Recovery Model II. As shown in Panel

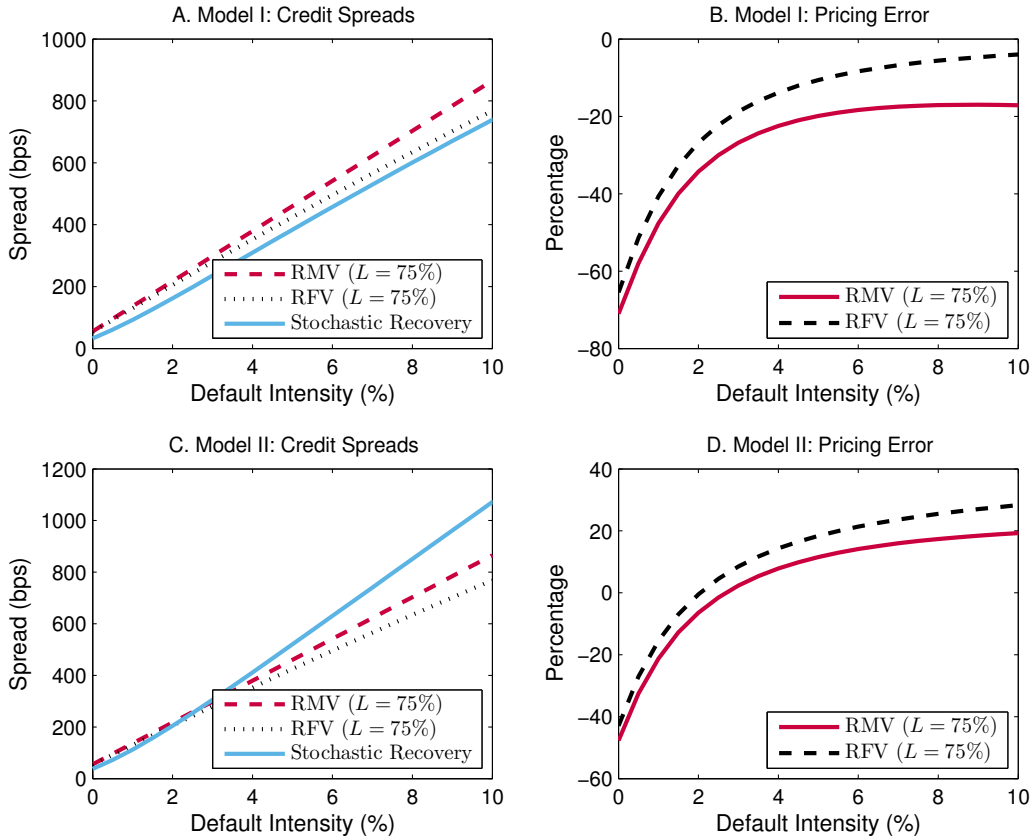


Figure 2: **Credit spreads for 5-year bonds with constant recovery and Cauchy recovery.** For different values of conditional default intensity, this figure plots the credit spreads of a 5-year zero-coupon defaultable bond, and the pricing errors of the RMV and RFV model with constant recovery rates relative to two versions of the stochastic recovery model. “RMV” stands for “recovery of market value”; “RFV” stands for “recovery of face value”.

C, the credit spreads generated by Model II are visibly more convex than the spreads from the RMV or RFV model with constant recovery rate. When default intensity is low, credit spreads are lower for the stochastic recovery model, because the conditional recovery rates are higher than under the constant recovery assumption. As aggregate default intensity rises, the size of the recovery risk premium increases, which lowers the risk-neutral recovery rates and explains the rapid rise in the spreads of Model II.

Panel D gives us a better sense of the size of pricing errors. When the aggregate default intensity is low, the recovery risk premium is small. Thus, the assumption of a 25% constant recovery rate during these times would be too “conservative”, which

makes the RMV model generate spreads that are too high. The pricing errors can be well over 20% for low values of λ . On the contrary, at times when the aggregate default rates are high, the assumption of 25% constant recovery rate becomes too optimistic. In fact, for $\lambda > 2.5\%$, the spreads in Model II exceed those in the RMV model. The pricing errors are over 10% for $\lambda > 4.5\%$, and can be as high as 20% when the default intensity reaches 10%.

There is another important message in [Figure 2](#). The curvature in the graphs of the spreads and pricing errors suggests that simply adjusting the constant recovery rate in the RMV model does not solve the mispricing problem. Changing the recovery rate amounts to (approximately) parallel-shifting the spreads or pricing errors, and will either exacerbate the underpricing for low λ or overpricing for high λ . These results suggest that it is important to dynamically account for the negative correlation between default intensity and recovery rate when pricing credit-sensitive securities.

4.2 Differences of Opinions

Models of heterogeneity of beliefs, or equivalently of preferences, can generate rich implications for trade and affect asset prices in equilibrium (see [Basak \(2005\)](#) for a recent survey). In studying such economies, aggregation often leads to difficulty in computing equilibrium outcomes. In this example, we illustrate the use of our main result in solving economies where there is heterogeneity among agents regarding beliefs (and higher order beliefs) about fundamentals.

4.2.1 General Setup

Suppose there are two agents (A, B) who possess heterogeneous beliefs. There is a state variable X_t which Agent A believes follows an affine jump-diffusion:

$$dX_t = \mu_t^A dt + \sigma_t^A dW_t^A + dZ_t^A, \quad (44)$$

where $\mu_t^A = K_0^A + K_1^A X_t$, $\sigma_t^A (\sigma_t^A)^\top = H_0^A + H_1^A \cdot X_t$, and jumps are believed to arrive with intensity $\lambda_t^A = \lambda_0^A + \lambda_1^A \cdot X_t$ and have distribution ν^A (with moment generating function ϕ^A). As elaborated in the examples below, the variable X_t encompass all uncertainty in the economy, including any time-variation in the heterogeneity of beliefs. For simplicity, we suppose that Agent A's beliefs are correct. The method is easily modified to the case where neither agent is correct.

Agent B has heterogeneous beliefs which we shall suppose are equivalent. A broad class¹² of such equivalent beliefs can be characterized as follows. There exists some vector a such that Agent B believes X follows an affine jump-diffusion satisfying

$$dX_t = \mu_t^B dt + \sigma_t^B dW_t^B + dZ_t^B, \quad (45)$$

where

1. $\mu_t^B = \mu_t^A + \sigma_t^A (\sigma_t^A)^\top a$
2. $\sigma_t^B = \sigma_t^A$
3. $d\nu^B/d\nu^A(Z) = e^{a \cdot Z}/E_{\nu^A}[e^{a \cdot Z}]$ or $\phi^B(c) = \phi^A(c + a)/\phi^A(a)$
4. $\lambda_t^B = \lambda_t^A \times E_{\nu^A}[e^{a \cdot Z}]$

This difference in beliefs generates a disagreement about not only the drifts of the state variables, but also the jump frequency and the distribution of jump size.¹³

This structure implies that the two beliefs define equivalent probability measures which may be related through the Radon-Nikodym derivative dP^B/dP^A :

$$\eta_t = E_t \left[\frac{dP^B}{dP^A} \right] = \exp \left(a \cdot X_t - \int_0^t \left(a \cdot \mu_s^A + \frac{1}{2} \|\sigma_s^A a\|^2 + \lambda_s^A (\phi_\nu(a) - 1) \right) ds \right). \quad (46)$$

The variable η_t expresses Agents B's differences in opinion in that when η_t is high, Agent B believes an event is more likely than Agent A believes. We refer to η_t as the *db-density* ('db' stands for "difference in beliefs") process, which differs from the density defining the risk-neutral measure.

While we specify the differences in beliefs exogenously, this does not preclude agents' beliefs from arising through Bayesian updating based on different information sets. For example, when the state variables and signals follow a joint Gaussian process, Bayesian updating can reduce to a difference of beliefs in the form of (46).

¹²More generally, we could consider beliefs of the form $e^{h(x_t) - \int_0^t e^{-h(x_s)} \mathcal{D}^1 e^{h(x_s)} ds}$. Provided the integral term remains tractable, the same analysis applies. Compare also the discussion of essentially affine difference of opinions.

¹³To be precise, as a process $Z^A = Z^B$ (i.e. the functions $Z^i : \Omega \times [0, \infty) \rightarrow \mathbb{R}$ are the same). Agents disagree about the probability measures on Ω .

Notice that the integral term in the exponent above follows an affine process. Thus, by redefining X to include the integral term and augmenting a accordingly, we have

$$\eta_t = e^{a \cdot X_t}. \quad (47)$$

We assume that the agents have time separable preferences:

$$U^i(c) = E_0^i \left[\int_0^\infty u^i(c_t, t) dt \right], \quad i = A, B. \quad (48)$$

Suppose also that

1. markets are complete;
2. log of aggregate consumption, $c_t = \log(C_t)$, is linear in X_t ($c_t = c \cdot X_t$);
3. agents are endowed with some fixed fraction ($\theta_A, \theta_B = 1 - \theta_A$) of aggregate consumption.

Let ξ_t denote the stochastic discount factor with respect to Agent A's beliefs. As in [Cox and Huang \(1989\)](#), we impose the lifetime budget constraint and equate state prices to marginal utilities to solve

$$u_c^A(C_t^A, t) = \zeta^A \xi_t, \quad (49)$$

$$u_c^B(C_t^B, t) = \zeta^B \eta_t^{-1} \xi_t, \quad (50)$$

where C_t^i is Agent i 's equilibrium consumption at time t and ζ^i is the Lagrange multiplier for Agent i 's budget constraint.

Market clearing then implies

$$C_t = (u_c^A)^{-1}(\zeta^A \xi_t) + (u_c^B)^{-1}(\zeta^B \eta_t^{-1} \xi_t), \quad (51)$$

which implies $\xi_t = h(c_t, \eta_t)$ for some h . With the additional assumption that $u^i(c, t) = e^{-\rho t} \frac{c^{1-\gamma}}{1-\gamma}$, this simplifies to

$$\xi_t = e^{-\rho t} \left[\left(\frac{1}{\zeta^A} \right)^{1/\gamma} + \left(\frac{\eta_t}{\zeta^B} \right)^{1/\gamma} \right]^\gamma C_t^{-\gamma}. \quad (52)$$

Using $g(x) = \left[\left(\frac{1}{\zeta^A} \right)^{1/\gamma} + \left(\frac{e^x}{\zeta^B} \right)^{1/\gamma} \right]^\gamma$ and $C_t = e^{c \cdot X_t}$, we finally have

$$\xi_t = e^{-\rho t} g(a \cdot X_t) e^{-\gamma c \cdot X_t}. \quad (53)$$

With the stochastic discount factor in this form, we may price any asset with *pl-linear* payoffs, such as bonds and dividend claims, using [Theorem 1](#).¹⁴ Our method also applies when the two agents have different risk aversion (γ_A and γ_B). In that case, we can still express $h(c_t, \eta_t)$ in the separable form as in (53), and proceed the same way.

In some cases, the mapping of a difference-of-opinion model to the standard setting (46) is not immediate, and requires a careful choice of the state variable X_t . For example, consider the setting where the agents believe that (de-trended) aggregate log consumption, c_t , follows an Ornstein-Uhlenbeck process:

$$dc_t = \kappa_A(\theta_A - c_t)dt + \sigma dW_t^A, \quad (54)$$

$$dc_t = \kappa_B(\theta_B - c_t)dt + \sigma dW_t^B. \quad (55)$$

In this case, the difference in beliefs cannot be expressed as in (46) directly. However, by considering an augmented state variable we can return to this form. The state variable $\langle c_t, c_t^2 \rangle$ follows the process

$$d \begin{bmatrix} c_t \\ c_t^2 \end{bmatrix} = \begin{bmatrix} \kappa_A(\theta_A - c_t) \\ 2c_t\kappa_A(\theta_A - c_t) + \frac{1}{2}\sigma^2 \end{bmatrix} dt + \begin{bmatrix} \sigma \\ 2c_t\sigma \end{bmatrix} dW_t^A. \quad (56)$$

Since the corresponding 2×2 conditional covariance matrix, $[\sigma, 2c_t\sigma]^\top [\sigma, 2c_t\sigma]$, is affine in $\langle c_t, c_t^2 \rangle$, it follows that $\langle c_t, c_t^2 \rangle$ is an affine process. Moreover, we return to our standard case since P^B is given by the change of measure as in (46) with

$$a = \sigma^{-2} \begin{bmatrix} \kappa_B\theta_B - \kappa_A\theta_A \\ \frac{1}{2}(\kappa_A - \kappa_B) \end{bmatrix}. \quad (57)$$

¹⁴The function g is not bounded and in fact does not even define a tempered function. Thus, our theory does not directly apply. One option is to write $g(x) = g_-(x)e^{-x} + g_+(x)e^{+x}$ where $g_\pm(x) = g(x)1_{\{\pm x < 0\}}e^{\mp x}$. Here g_\pm are bounded functions whose Fourier transforms can be computed in terms of incomplete Beta functions. Another option is to write $g(x) = g(x)^{\lceil \gamma \rceil / \gamma} g(x)^{-\lceil \gamma \rceil / \gamma + 1}$. In this case, the first functional is *pl-linear* and the second is bounded with Fourier transform known in terms of Beta functions.

More generally, we can have the case where each agent believes that the state of the economy is summarized by the N -dimensional Gaussian state variables, X_t , and each agent believes that X_t satisfies the stochastic differential equation $dX_t = (K_0^i + K_1^i X_t)dt + \sqrt{H_0}dW_t^i$. Again by considering an augmented state variable of the form $\hat{X}_t = \langle X_t, \text{vech}(X_t X_t^\top) \rangle$ we can return to our standard setting.¹⁵ Such techniques are common in the term structure literature with respect to affine and quadratic term structure models. The procedure generalizes to accommodate models with stochastic volatility ($A_M(N)$ in the parlance of [Dai and Singleton \(2000\)](#)). Following [Duffee \(2002\)](#), we refer to this as *essentially affine difference of beliefs*.

An alternative characterization is to consider the “market price of belief risk”, λ_t , in analogy to the usual market price of risk. By defining

$$\lambda_t = \sqrt{H_0^{-1}}(\mu_t^B - \mu_t^A), \quad (58)$$

$$\eta_t = e^{-\int_0^t \lambda_s dW_s^A - \frac{1}{2} \int_0^t \|\lambda_s\|^2 ds}. \quad (59)$$

When η_t is exponential affine in X_t , this defines an appropriate Radon-Nikodym derivative for our setting.

4.2.2 Special Cases

The framework above can accommodate a wide range of specifications with heterogeneity of beliefs regarding expected changes in fundamentals, likelihood of jumps, distribution of jumps, and divergence in higher order beliefs. We now provide some examples.

Disagreement about stochastic growth rates. This is the model studied in [Dumas, Kurshev, and Uppal \(2009\)](#), hereafter DKU. In their model, there is a single dividend process C_t with time-varying growth rate, but agents A and B have different beliefs regarding the growth rate of the tree, \hat{f}_t^A and \hat{f}_t^B , and $\hat{g}_t = \hat{f}_t^B - \hat{f}_t^A$ represents the amount of disagreement between B and A.

This model can be mapped into the essentially affine difference in beliefs specification, and our results can simplify the calculations for the most general model that

¹⁵For a square matrix M , vech denotes the lower triangular entries written as a vector. Usually, only part of the elements in the extended state vector is needed to maintain the Markov structure.

they consider. First, under Agent B's probability measure,

$$d \begin{bmatrix} c_t \\ \hat{f}_t^B \\ \hat{g}_t \end{bmatrix} = \begin{bmatrix} \hat{f}_t^B - \frac{1}{2}\sigma_c^2 \\ \kappa(\bar{f} - \hat{f}_t^B) \\ -\psi\hat{g}_t \end{bmatrix} dt + \begin{bmatrix} \sigma_c & 0 \\ \frac{\gamma_B}{\sigma_c} & 0 \\ \sigma_{\hat{g},c} & \sigma_{\hat{g},s} \end{bmatrix} dW_t^B. \quad (60)$$

Next, in order to map the model to our standard setting, we define the augmented state variable as $X_t = \langle c_t, \log \eta_t, \hat{f}_t^B, \hat{g}_t, \hat{g}_t^2 \rangle$, where η_t gives the density process: $\eta_t = E_t[dP^A/dP^B]$. The dynamics of X_t are given by the stochastic differential equation:

$$dX_t = (K_0 + K_1 X_t)dt + \Sigma_t dW_t^B,$$

where

$$K_0 = \begin{bmatrix} -\frac{1}{2}\sigma_c^2 \\ 0 \\ \kappa\bar{f} \\ 0 \\ \sigma_{\hat{g},c}^2 + \sigma_{\hat{g},s}^2 \end{bmatrix}, \quad K_1 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2\sigma_c^2} \\ 0 & 0 & -\kappa & 0 & 0 \\ 0 & 0 & 0 & -\psi & 0 \\ 0 & 0 & 0 & 0 & -2\psi \end{bmatrix}, \quad \Sigma_t = \begin{bmatrix} \sigma_c & 0 \\ -\hat{g}_t/\sigma_c & 0 \\ \gamma_B/\sigma_c & 0 \\ \sigma_{\hat{g},c} & \sigma_{\hat{g},s} \\ 2\sigma_{\hat{g},c}\hat{g}_t & 2\sigma_{\hat{g},s}\hat{g}_t \end{bmatrix}.$$

It is easy to check that the local conditional variance of X_t , $\Sigma_t \Sigma_t^\top$, is affine in X_t so this represents an affine process.¹⁶ Then, it is immediate that η_t takes the form of (46) with $a = \langle 0, 1, 0, 0, 0 \rangle$.

DKU show that in their setting a number of equity and fixed income security prices take the form $E_0[e^{\alpha \cdot X_t} g(\beta \cdot X_t)]$ where $g(x) = (1 - e^{ax})^b$ for some (α, β, a, b) . They use two methods to compute this moment. First, when $b \in \mathbb{N}$, g can be expanded directly and reduced to log-linear functionals. Then the moments can be computed by well-known methods. For more general cases, they compute the moment in two steps: first recover the forward density of $\beta \cdot X$ through a Fourier inversion of the conditional characteristic function, and then evaluate the expectation using the density. The formula (A58-A61) in DKU is essentially

$$E_0[e^{\alpha \cdot X_t} g(\beta \cdot X_t)] = \frac{1}{2\pi} \int_{b \in \mathbb{R}} \hat{g}(b) \int_{s \in \mathbb{R}} e^{ibs} E_0[e^{(\alpha - is\beta) \cdot X_t}] ds db. \quad (61)$$

¹⁶DKU exploit the fact that in this particular case the ODE determining the conditional characteristic function for some variables can be computed in closed form by standard methods. However, in general there is little additional complication to solve the usual ODE by standard numerical methods.

This formula requires a *double integral*, thus increasing the dimensionality of the problem. As Theorem 1 shows, our generalized transform method will only require a *single integral* to compute this moment. If we consider the generalization $g(\beta_1 \cdot X_t, \beta_2 \cdot X_t)$, the trade-off becomes a somewhat tractable 2-dimensional integral with our method versus a highly intractable 4-dimensional integral by using an extension of the DPS method.

Disagreement about volatility. Suppose that dividends have stochastic volatility. Under Agent A's beliefs:

$$d \begin{bmatrix} c_t \\ V_t \end{bmatrix} = \begin{bmatrix} \bar{g} \\ -\kappa_V V_t \end{bmatrix} dt + \sqrt{\begin{bmatrix} \sigma_d & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \sigma_{cV} V_t & 0 \\ 0 & \sigma_{VV} V_t \end{bmatrix}} dW_t^A. \quad (62)$$

Here σ_d is the lowest conditional variance of log dividends, while V_t represents the degree to which volatility is above the lowest level.

Agent B disagrees about the dynamics of volatility. According to his beliefs:

$$d \begin{bmatrix} c_t \\ V_t \end{bmatrix} = \begin{bmatrix} \bar{g} \\ -(\kappa_V - b)V_t \end{bmatrix} dt + \sqrt{\begin{bmatrix} \sigma_d & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \sigma_{cV} V_t & 0 \\ 0 & \sigma_{VV} V_t \end{bmatrix}} dW_t^B. \quad (63)$$

For example, when $b > 0$, Agent B believe that volatility mean reverts more slowly. Using $a = \langle 0, b/\sigma_{VV}^2 \rangle$ we get the *db-density* as in (46).

Disagreement about momentum. Consider a model with stochastic growth in consumption. Let c_t be the log consumption, g_t be the expected growth rate. Also, let e_t be an exponential weighted moving average of past growth rates:

$$e_t = \int_{-\infty}^t e^{-b(t-s)} g_s ds. \quad (64)$$

Agent A correctly believes that the expected growth rate of log consumption is g_t . Under her beliefs:

$$d \begin{bmatrix} c_t \\ g_t \\ e_t \end{bmatrix} = \begin{bmatrix} g_t \\ \kappa(\bar{g} - g_t) \\ g_t - be_t \end{bmatrix} dt + \begin{bmatrix} \sigma_c & 0 \\ 0 & \sigma_g \\ 0 & 0 \end{bmatrix} dW_t^A. \quad (65)$$

Agent B believes that growth is due to two components: (1) a mean-reverting component, g_t and (2) a counteracting momentum component through e_t .

$$d \begin{bmatrix} c_t \\ g_t \\ e_t \end{bmatrix} = \begin{bmatrix} g_t + ce_t \\ \kappa(\bar{g} - g_t) \\ g_t - be_t \end{bmatrix} dt + \begin{bmatrix} \sigma_c & 0 \\ 0 & \sigma_g \\ 0 & 0 \end{bmatrix} dW_t^B. \quad (66)$$

Fixing the past, for large enough deviations from the steady-state, the mean-reverting component will dominate. However, for small deviation from the steady state, Agent B will believe that past deviations from the steady state lead to larger future deviations from the state steady. In this way we can view Agent B as possessing a conservatism or “law of small numbers” bias.

This example represents a special case of the essentially affine difference of beliefs.

Disagreement about higher order beliefs. Heterogeneity in higher order beliefs can affect asset prices as well. We can inductively proceed in defining beliefs:

\hat{g}_t^i = Agent i 's beliefs about the growth rate of consumption

\hat{g}_t^{ij} = Agent i 's beliefs about Agent j 's belief about the growth rate of consumption

We can consider the state variable $X_t = [c_t, \hat{g}_t^A, \hat{g}_t^B, \hat{g}_t^{AB}, \hat{g}_t^{BA}]$. Suppose that X_t follows a Gaussian process under both agents beliefs. Agent A's beliefs are such that

$$d \begin{bmatrix} c_t \\ \hat{g}_t^A \\ \hat{g}_t^B \\ \hat{g}_t^{AB} \\ \hat{g}_t^{BA} \end{bmatrix} dt = \begin{bmatrix} \hat{g}_t^A \\ \kappa_A(\theta - \hat{g}_t^A) \\ \kappa_B(\theta - \hat{g}_t^B) \\ \kappa_{AB}(\hat{g}_t^B - \hat{g}_t^{AB}) \\ \kappa_{BA}(\hat{g}_t^A - \hat{g}_t^{BA}) \end{bmatrix} dt + \Sigma dW_t^A. \quad (67)$$

Here, the fourth and fifth components of the drift say that Agent A believes that the higher order beliefs (both his beliefs about Agent B and Agent B's beliefs about him) are correct in the long run, but may have short run deviations.

Again, this model represent a special case of the essentially affine disagreement.

Disagreement about the likelihood of disasters. Suppose that log consumption, c_t , has constant growth with IID innovations with time-varying probability, λ_t , of rare

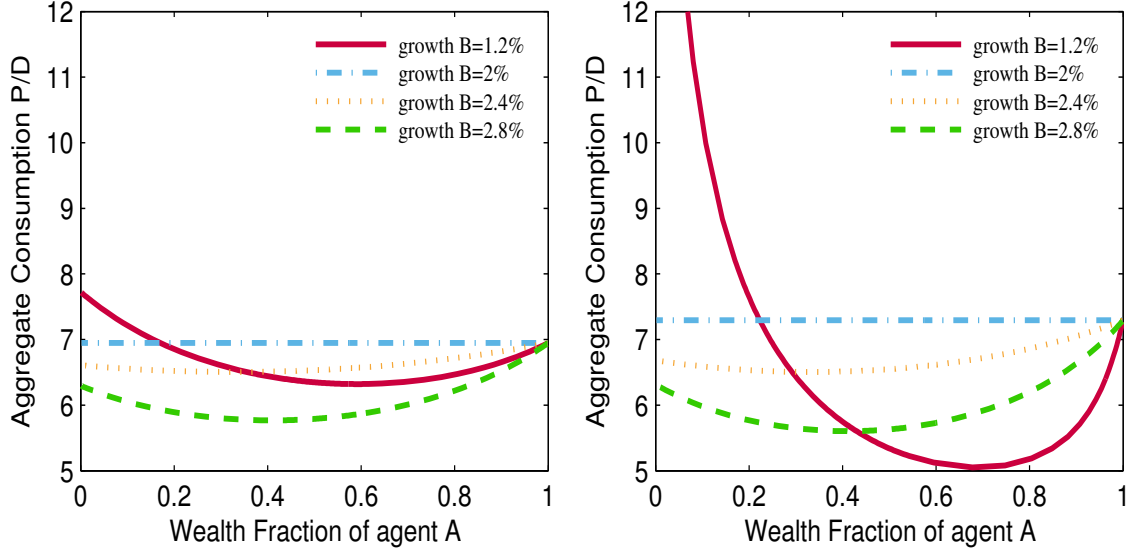


Figure 3: **Price-Dividend Ratios with Heterogeneity in Beliefs.** These figures plots the Price-Dividend ratio in an economy where agents hold different beliefs regarding growth rates and disaster probabilities.

disaster. Let $X_t = [c_t, \lambda_t]$. Under Agent A's beliefs,

$$dX_t = \begin{bmatrix} g_A \\ -\kappa_\lambda \lambda_t \end{bmatrix} dt + \begin{bmatrix} \sigma_c & 0 \\ 0 & \sigma_\lambda \sqrt{\lambda} \end{bmatrix} dW_t^A + dZ_t^A, \quad (68)$$

where Z_t^A are jumps in c_t which occur with intensity $\lambda_0 + \lambda_t$ and distribution ν . Suppose that Agent B's beliefs are specified by the db-density of form (46) with $a = \langle b, 0 \rangle$. Then, Agent B's beliefs will be

$$dX_t = \begin{bmatrix} g_A + b\sigma_c^2 \\ -\kappa_\lambda \lambda_t \end{bmatrix} dt + \begin{bmatrix} \sigma_c & 0 \\ 0 & \sigma_\lambda \sqrt{\lambda} \end{bmatrix} dW_t^B + dZ_t^B, \quad (69)$$

where jumps arrive with intensity $\lambda_t^B = E_{\nu^A}[e^{a \cdot Z}](\lambda_0 + \lambda_t)$ and have distribution ν^B with Radon-Nikodym derivative $d\nu^B/d\nu^A(Z) = e^{a \cdot Z}/E_{\nu^A}[e^{a \cdot Z}]$.

In this sense, Agent B is more optimistic about the future growth both in terms of (1) higher expected growth rates, (2) lower likelihood of disasters, (3) less severe losses conditional on there being a disaster.

For illustration, we consider the case of disagreement about the likelihood of

disasters. The model is calibrated as follows:

- $g_A = 2\%$, $\sigma_c = 1.5\%$, $\kappa_\lambda = 0.1$, $\lambda_0 = 1.5\%$, $\sigma_\lambda^\infty = 1\%$, (the volatility of the stationary distribution of λ), $\gamma = 2$. Each jump causes a 8% drop in consumption ($Z = -8\%$).
- Agent B's beliefs are generated with $a = \langle b, 0 \rangle$ where b is chosen so that the agent is either more pessimistic or optimistic (both with regard to growth rates and the likelihood of disasters).
- Lagrange multiplier for agent A, ζ^A , is adjusted to vary the fraction of wealth endowed to agent A. (ζ^B is normalized to 1.)

We vary the amount of disagreement between the two agents by changing b . A more negative value for b simultaneously lowers Agent B's perceived growth rate g_B , and increases his perceived likelihood of a disaster. More specifically, the growth rate of consumption (excluding jumps) under Agent B's belief is $g_B = g_A + b\sigma_c^2$, and the jump intensity under B's belief is e^{bZ} times the intensity under A's belief.

Figure 3 plots the price-dividend ratio of the aggregate consumption claim. Agent A believes the expected growth rate of consumption is $g_A = 2\%$. Agent B can be pessimistic or optimistic relative to Agent A. The horizontal axis gives the fraction of total wealth in the economy owned by Agent A at $t = 0$. The left panel considers the special case without disaster risk ($Z = 0$), and the right panel considers the case with disasters. First, in an economy where only Agent A lives, the price-dividend ratio is 6.9 (the dashed line). When we introduce Agent B, who is more pessimistic than A (B believes in lower expected growth rate and higher chances of disasters), the equilibrium price of the consumption claim drops. Interestingly, as the wealth of Agent B (A) increases (decreases), the price first decreases and then increases. Finally, when Agent B owns all the wealth, the price-dividend ratio rises to 7.8. The results are qualitatively similar when Agent B is more optimistic.

The eventual increase of asset prices with the wealth of the pessimistic agent is due to the low elasticity of intertemporal substitution of a CRRA-utility agent with $\gamma > 1$. The agent wants to smooth consumption by saving more, which lowers the interest rate and raises the price of consumption claims. The reason for the initial decline of asset prices is as follows. Consider Agent A, who has more optimistic beliefs. He sells Agent B his share on aggregate consumption in the "bad states" (where aggregate

consumption is low) in exchange for a bigger share on aggregate consumption in the “good states” (where aggregate consumption is high). In equilibrium, his consumption has higher expected growth rate and volatility than the aggregate endowment. As a result, the interest rate rises, and so does the Sharpe ratio under his beliefs. For large enough disagreement in beliefs, this effect will dominate, causing prices to fall.

When there is disagreement about the probability of disasters, the impact of heterogeneous beliefs on prices becomes more striking. As shown in the right panel, when Agent B is more pessimistic, the equilibrium price of consumption claims falls by over 14% even when Agent B only owns 10% of the total wealth in the economy. This result is consistent with the findings of [Kogan, Ross, Wang, and Westerfield \(2006\)](#) that irrational traders can still have nontrivial price impacts when their wealth is small.

5 Concluding Remarks

We extend the transform analysis in [Duffie, Pan, and Singleton \(2000\)](#) to compute a general class of nonlinear moments for affine jump-diffusions. Through a Fourier decomposition of the nonlinear moments, we can directly utilize the properties of the conditional characteristic functions for affine processes and compute the moments analytically. By not resorting to an intermediate computation of the (forward) density, this method greatly reduces the dimensionality of such problems, allowing for tractability in a wide range of economic applications.

We demonstrate the power of this method with examples from several areas, including option pricing, term structure, GMM estimation, and an equilibrium asset pricing model both with multiple agents and multiple goods. Underlying all of these examples are the rich dynamics provided by affine processes, allowing for time-varying conditional means and variances as well as jumps occurring with stochastic intensity. We also illustrate the application of the generalized transform method in two in-depth examples: a model of defaultable bond pricing, where the recovery rates are conditionally correlated with default intensities and a general class of difference-of-opinion models.

Appendix

A Proof of Theorem 1

Throughout, we maintain the following assumptions:

Assumption 1: In the terminology of DPS, (Θ, α, β) is well-behaved at (s, T) for all $s \in \mathbb{R}$. That is,

- (a) $E\left(\int_0^T |\gamma_t| dt\right) < \infty$ where $\gamma_t = \Psi_t(\phi(B(T-t)) - 1)(\lambda_0 + \lambda_1(X_t))$
- (b) $E[(\int_0^T \|\eta_t\|^2 dt)] < \infty$ where $\|\eta_t\|^2 = \Psi_t^2 B(T-t)^\top (H_0 + H_1 \cdot X_t) B(T-t)$
- (c) $E[|\Psi_T|] < \infty$

where $\Psi_t = e^{-\int_0^t r_s ds} e^{A(T-t) + B(T-t) \cdot X_t}$ and A, B solve the ODE given in (8-9).

Assumption 2: The measure F defined by its Radon-Nikodym derivative,

$$\frac{dF}{dP} = \frac{e^{-\int_0^T r_\tau d\tau} e^{\alpha \cdot X_T}}{E_0[e^{-\int_0^T r_\tau d\tau} e^{\alpha \cdot X_T}]}, \quad (70)$$

is such that the density of $\beta \cdot X_T$ under F is a Schwartz function. In particular, the density of $\beta \cdot X_T$ is smooth and declines faster than any polynomial under F .

Proposition 1 of DPS gives conditions under which Assumption 1 holds. These are integrability conditions which imply that, for every s , the local martingale

$$E_t[e^{-\int_t^T r_\tau d\tau + \alpha + is\beta}] e^{-A_{T-t} - B_{T-t} \cdot X_t}$$

is in fact a martingale.

Assumption 2 is analogous to (2.11) of DPS. However, we require a somewhat stronger assumption to directly apply our theory. This assumption can typically be shown to hold by verifying that the moment generating function (under F) is finite in a neighborhood of 0. See [Duffie, Filipovic, and Schachermayer \(2003\)](#).

We now prove [Theorem 1](#). Suppose now that Assumptions 1 and 2 hold. Then,

$$\begin{aligned} H &= E_0[e^{-\int_0^T r_\tau d\tau} e^{\alpha \cdot X_T} g(\beta \cdot X_T)] \\ &= F_0 E_0^F[g(\beta \cdot X_T)] \\ &= F_0 \int g(b) f_{\beta \cdot X_T}^F(b) db \\ &= F_0 \langle g, f_{\beta \cdot X_T}^F \rangle. \end{aligned}$$

In the last equation, we interpret $g \in \mathcal{S}'$. By Assumption 2, $f_{\beta \cdot X_T}^F \in \mathcal{S}$, and so $\hat{f}_{\beta \cdot X_T}^F \in \mathcal{S}$ also. Thus Fourier inversion holds and $(\hat{f}_{\beta \cdot X_T}^F)^\check{=} = \frac{1}{2\pi} f_{\beta \cdot X_T}^F$ (see Corollary 8.28 in [Folland \(1984\)](#).)¹⁷ Applying this,

$$\begin{aligned} H &= \frac{1}{2\pi} F_0 \langle g, (\hat{f}_{\beta \cdot X_T}^F)^\check{=} \rangle \\ &= \frac{1}{2\pi} F_0 \langle \hat{g}, \check{f}_{\beta \cdot X_T}^F \rangle \\ &= \frac{1}{2\pi} \langle \hat{g}, F_0 \check{f}_{\beta \cdot X_T}^F \rangle \\ &= \frac{1}{2\pi} \langle \hat{g}, G(\alpha - \cdot \beta i) \rangle. \end{aligned}$$

The second step holds by the definition of the Fourier transform of a tempered distribution and the last step hold by Assumption 1. This is the desired result. \square

B Proof of Proposition 1

In analogy to [Duffie, Pan, and Singleton \(2000\)](#) and [Pan \(2002\)](#), define

$$G(\alpha_0; v, n|x, t) = e^{A_t + B_t \cdot x} \sum_{|\xi|=n} \binom{n}{\xi} L(x)^\xi \quad (71)$$

where $L(x)$ is the n -dimensional vector whose i th coordinate is $(\partial_i A + \partial_i B \cdot x)^{1/i}$, ξ is a n -dimensional multi-index, and $(\partial_i A, \partial_i B)_i$ satisfies the ODE

$$\dot{B} = K_1^\top B + \frac{1}{2} B^\top H_1 B - \rho_1 + \lambda_1(\phi(B) - 1) \quad B(0) = \alpha_0 \quad (72)$$

$$\dot{A} = K_0^\top B + \frac{1}{2} B^\top H_0 B - \rho_0 + \lambda_0(\phi(B) - 1) \quad A(0) = 0 \quad (73)$$

$$\partial_1 \dot{B} = K_1^\top \partial_1 B + \partial_1 B^\top H_1 B + \lambda_1 \nabla \phi(B) \cdot \partial_1 B \quad \partial_1 B(0) = v \quad (74)$$

$$\partial_1 \dot{A} = K_0^\top \partial_1 B + \partial_1 B^\top H_0 B + \lambda_0 \nabla \phi(B) \cdot \partial_1 B \quad \partial_1 A(0) = 0 \quad (75)$$

¹⁷We use the convention that for $f \in L^1$, $\hat{f}(s) = \int e^{-ix \cdot s} f(x) dx$, $\check{f}(x) = \int e^{ix \cdot s} f(s) ds$ and for $g \in \mathcal{S}'$, $\langle \hat{g}, f \rangle \equiv \langle g, \check{f} \rangle$.

and for $2 \leq m \leq n$, $(\partial_m B, \partial_m A)$ satisfy

$$\partial_m \dot{B} = K_1^\top \partial_1 B + \frac{1}{2} \sum_{i=0}^m \binom{m}{i} \partial_i B^\top H_1 \partial_{m-i} B + \partial_{m-1} (\lambda_1 \nabla \phi(B) \cdot \partial_1 B) \quad \partial_m B(0) = 0 \quad (76)$$

$$\partial_m \dot{A} = K_0^\top \partial_1 B + \frac{1}{2} \sum_{i=0}^m \binom{m}{i} \partial_i B^\top H_1 \partial_{m-i} B + \partial_{m-1} (\lambda_0 \nabla \phi(B) \cdot \partial_1 B) \quad \partial_m A(0) = 0 \quad (77)$$

We strengthen Assumptions 1 and 2 as follows:

1. **Assumption 1'**: The moment generating function, $\phi \in C^N(D_0)$ where D_0 is an open set containing the image of the solutions to (8) for any initial condition of the form $\alpha_0 = \alpha + is\beta$ for any $s \in \mathbb{R}$. Additionally, for any such a initial condition:

(a) $E\left(\int_0^T |\gamma_t| dt\right) < \infty$ where

$$\gamma_t = \lambda_t E_\nu[\Psi_t^n(i_t, X_t + Z) - \Psi_t^n(i_t, X_t)]$$

and $\Psi_t^n(i, x) = e^{-iG(\alpha, \nu, n|x, T-t)}$ and $i_t = \int_0^t r_s ds$.

(b) $E[(\int_0^T \|\eta_t\|^2 dt)] < \infty$ where

$$\|\eta_t\|^2 = \nabla_x \Psi_t^n(i_t, X_t)^\top (H_0 + H_1 \cdot X_t) \nabla_x \Psi_t^n(i_t, X_t)$$

(c) $E[|\Psi_T(i_T, X_T)|] < \infty$

2. **Assumption 2'**: The measure F defined by its Radon-Nikodym derivative,

$$\frac{dF}{dP} = \frac{e^{-\int_0^T r_\tau d\tau} e^{\alpha \cdot X_T} (\nu \cdot X_T)^n}{E_0[e^{-\int_0^T r_\tau d\tau} e^{\alpha \cdot X_T} (\nu \cdot X_T)^n]}, \quad (78)$$

is such that the density of $\beta \cdot X_T$ under F is a Schwartz function.

Given Assumption 1' and Assumption 2' hold, the proof follows as before.

C Deriving the stochastic discount factor in Section 3.5

From the first order conditions, we obtain

$$[C_{A,1}^\alpha + \bar{\delta}C_{A,2}^\alpha]^{\frac{1-\gamma}{\alpha}-1} C_{A,1}^{\alpha-1} = \lambda [\bar{\delta}C_{B,1}^\alpha + C_{B,2}^\alpha]^{\frac{1-\gamma}{\alpha}-1} C_{B,1}^{\alpha-1} \bar{\delta}, \quad (79)$$

$$[C_{A,1}^\alpha + \bar{\delta}C_{A,2}^\alpha]^{\frac{1-\gamma}{\alpha}-1} \bar{\delta}C_{A,2}^{\alpha-1} = \lambda [\bar{\delta}C_{B,1}^\alpha + C_{B,2}^\alpha]^{\frac{1-\gamma}{\alpha}-1} C_{B,2}^{\alpha-1}, \quad (80)$$

where we have renormalized the home bias $\bar{\delta} = \frac{\delta}{1-\delta}$. This implies

$$\left(\frac{C_{A,2}}{C_{A,1}}\right)^{\alpha-1} = \bar{\delta}^2 \left(\frac{C_{B,2}}{C_{B,1}}\right)^{\alpha-1}. \quad (81)$$

From (33a–33b) and (79–80), we can solve for the equilibrium consumption $C_{A,1}$ and $C_{A,2}$ in terms of D_1 and D_2 . Letting $f_{i,j} = C_{i,j}/D_j$, Agent i 's fraction of the consumption of good j . Then (81) implies $f_{A,2}/(1-f_{A,2}) = \bar{\delta}^{2/(\alpha-1)} \times f_{A,1}/(1-f_{A,1})$ or

$$f_{A,2} = \frac{f_{A,1}}{\bar{\delta}^{\frac{2}{\alpha-1}} + (1-\bar{\delta}^{\frac{2}{\alpha-1}})f_{A,1}} \quad (82)$$

Thus the equilibrium fraction of consumption of agent A of good 2 is a monotonic function of the fraction of consumption of good 1. Additionally, substituting into (79), we obtain

$$\left[f_{A,1}^\alpha + \bar{\delta} \left(\frac{D_2}{D_1}\right)^\alpha f_{A,2}^\alpha\right]^{\frac{1-\gamma}{\alpha}-1} f_{A,1}^{\alpha-1} = \lambda \left[\bar{\delta}(1-f_{A,1})^\alpha + \left(\frac{D_1}{D_2}\right)^\alpha (1-f_{A,2})^\alpha\right]^{\frac{1-\gamma}{\alpha}-1} (1-f_{A,1})^{\alpha-1} \bar{\delta}. \quad (83)$$

Thus the fraction of consumption of each good for Agent A is a function of the (exogenous) relative supply of each good, $R_t = D_{2,t}/D_{1,t}$. Moreover, using (82) and monotonicity we see that the equilibrium consumption fraction $f_{A,1}$ is in some interval $[\underline{f}_A, \bar{f}_A] \subset (0, 1)$. Let us define $\hat{f}_{A,1}(\frac{D_2}{D_1})$ to be the implicit solution to (82–83). Similarly, we can define implicit functions $\hat{f}_{A,2}, \hat{f}_{B,1}, \hat{f}_{B,2}$ which give the equilibrium consumption fractions of each good in terms of the relative supplies of the goods.

Thus, the stochastic discount factor for Agent A is

$$\begin{aligned}
\xi_A(t) &= e^{-\rho t} \frac{\partial u}{\partial C_{A,1}} \\
&= e^{-\rho t} \left[\left(D_1 \times \hat{f}_{A,1} \left(\frac{D_2}{D_1} \right) \right)^\alpha + \left(D_2 \times \hat{f}_{A,2} \left(\frac{D_2}{D_1} \right) \right)^\alpha \right]^{\frac{1-\gamma-\alpha}{\alpha}} \left(D_1 \times \hat{f}_{A,1} \left(\frac{D_2}{D_1} \right) \right)^{\alpha-1} \\
&= e^{-\rho t} \left[R^{-\alpha} \left(\hat{f}_{A,1}(R) \right)^\alpha + R^\alpha \left(\hat{f}_{A,2}(R) \right)^\alpha \right]^{\frac{1-\gamma-\alpha}{\alpha}} \left(\hat{f}_{A,1}(R) \right)^{\alpha-1} D_1^{\frac{-1-\gamma+\alpha}{2}} D_2^{\frac{1-\gamma-\alpha}{2}}.
\end{aligned} \tag{84}$$

where $R = D_2/D_1$. Define g to be the implicit function

$$g(R) = \left[R^{-\alpha} \left(\hat{f}_{A,1}(R) \right)^\alpha + R^\alpha \left(\hat{f}_{A,2}(R) \right)^\alpha \right]^{\frac{1-\gamma-\alpha}{\alpha}} \left(\hat{f}_{A,1}(R) \right)^{\alpha-1},$$

we then get the stochastic discount factor in (34).

References

- Edward I. Altman, Brooks Brady, Andrea Resti, and Andrea Sironi. The link between default and recovery rates: Theory, empirical evidence, and implications. *Journal of Business*, 78(6):2203–2228, 2005.
- Gurdip Bakshi and Dilip Madan. Spanning and derivative-security valuation. *Journal of Financial Economics*, 55(2):205–238, 2000.
- Gurdip Bakshi, Charles Cao, and Zhiwu Chen. Empirical performance of alternative option pricing models. *Journal of Finance*, 52(5):2003–49, 1997.
- Gurdip Bakshi, Dilip Madan, and Frank Zhang. Understanding the role of recovery in default risk models: Empirical comparisons and implied recovery rates. Working paper, 2006.
- Suleyman Basak. Asset prices with heterogenous beliefs. *Journal of Banking & Finance*, 29:28492881, 2005.
- David S. Bates. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of Financial Studies*, 9(1):69–107, Spring 1996.
- Peter Carr and Dilip B. Madan. Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2:61–73, 1999.
- Hui Chen. Macroeconomic conditiosn and the puzzles of credit spreads and capital structure. Working Paper, MIT, 2008.
- Ren-Raw Chen and Louis Scott. Interest rate options in multifactor cox-ingersoll-ross models of the term structure. *Journal of Derivatives*, 3:53–72, 1995.
- John H. Cochrane, Francis Longstaff, and Pedro Santa-Clara. Two trees. *Review of Financial Studies*, 21(1):347–385, 2008.
- Pierre Collin-Dufresne, Robert Goldstein, and Julien Hugonnier. A general formula for valuing defaultable securities. *Econometrica*, 72(5):1377–1407, 2004.
- John C. Cox and Chifu Huang. Optimum consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory*, 49:33–83, 1989.

- Qiang Dai and Kenneth J. Singleton. Specification analysis of affine term structure models. *Journal of Finance*, 55(5):1943–1978, October 2000.
- Gregory R. Duffee. The relation between treasury yields and corporate bond yield spreads. *Journal of Finance*, 53:2225–2242, 1998.
- Gregory R. Duffee. Term premia and interest rate forecasts in affine models. *Journal of Finance*, 57(1):405–443, February 2002.
- Gregory R Duffee. Estimating the price of default risk. *Review of Financial Studies*, 12(1):197–226, 1999.
- J. Darrell Duffie. Credit risk modeling with affine processes. *Journal of Banking and Finance*, 29:2751–2802, 2005.
- J. Darrell Duffie and Kenneth J. Singleton. Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12:687–720, 1999.
- J. Darrell Duffie, Mark Schroder, and Costis Skiadas. Recursive valuation of defaultable securities and the timing of resolution of uncertainty. *Annals of Applied Probability*, 6(4):1075–1090, 1996.
- J. Darrell Duffie, Jun Pan, and Kenneth Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, November 2000.
- J. Darrell Duffie, Damir Filipovic, and Walter Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13(3):984–1053, 2003.
- Bernard Dumas, Alexander Kurshev, and Raman Uppal. Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility. *Journal of Finance*, *Forthcoming*, 2009.
- Gerald Folland. *Real Analysis: Modern Techniques and Their Applications*. Wiley-Interscience, 1984.
- Steve L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343, 1993.

- Scott Joslin. Pricing and hedging volatility in fixed income markets. Working paper, MIT, 2009.
- Leonid Kogan, Stephen A. Ross, Jiang Wang, and Mark M. Westerfield. The price impact and survival of irrational traders. *Journal of Finance*, 61(1):195–229, 2006.
- Francis Longstaff and Eduardo Schwartz. A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance*, 50:789–819, 1995.
- Robert E. Lucas, Jr. Asset prices in an exchange economy. *Econometrica*, 46:1426–14, 1978.
- Ian Martin. The lucas orchard. 2008. Working Paper.
- Jun Pan. The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics*, 63(1):3–50, January 2002.
- Jun Pan and Kenneth J. Singleton. Default and recovery implicit in the term structure of sovereign "cds" spreads. *Journal of Finance*, 63(5):2345–2384, 2008.
- Anna Pavlova and Roberto Rigobon. Asset prices and exchange rates. *Review of Financial Studies*, 20(4):1139–1181, 2007.
- Monika Piazzesi. Bond yields and the federal reserve. *Journal of Political Economy*, 113(2):311–344, April 2005.
- Monika Piazzesi, Martin Schneider, and Selale Tuzel. Housing, consumption and asset pricing. *Journal of Financial Economics*, 83(3):531–569, 2007.
- Philip E. Protter. *Stochastic Integration and Differential Equations*. Springer-Verlag, New York, NY, 2nd edition, 2003.
- Kenneth J. Singleton. Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics*, 102:111–141, 2001.
- John B. Taylor. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy*, 39:195–214, 1993.