Generalized Transform Analysis of Affine Processes and Applications in Finance

Hui Chen
MIT Sloan School of Management and NBER

Scott Joslin
USC Marshall School of Business

Nonlinearity is an important consideration in many problems of finance and economics, such as pricing securities and solving equilibrium models. This article provides analytical treatment of a general class of nonlinear transforms for processes with tractable conditional characteristic functions. We extend existing results on characteristic function-based transforms to a substantially wider class of nonlinear functions while maintaining low dimensionality by avoiding the need to compute the density function. We illustrate the applications of the generalized transform in pricing defaultable bonds with stochastic recovery. We also use the method to analytically solve a class of general equilibrium models with multiple goods and apply this model to study the effects of time-varying labor income risk on the equity premium. (*JEL* E44, G12)

In this article, we provide analytical treatment of a class of transforms for processes with tractable characteristic functions. These transforms bring analytical and computational tractability to a large class of nonlinear moments, and can be applied in option pricing, structural estimation, or solving equilibrium asset pricing models. We demonstrate the utility of our method with two examples. One example is on pricing defaultable bonds with stochastic recovery; the other example is on solving a general equilibrium model with stochastic labor income risk.

Consider a state variable X_t with transition dynamics under a certain probability measure summarized by a tractable conditional characteristic function. Many popular stochastic processes in economics have simple characteristic functions, such as the affine jump-diffusions, Lévy processes,

We thank Peter Feldhutter, Damir Filipovic, Xavier Gabaix, Chris Jones, Martin Lettau, Ian Martin, Jun Pan, Tano Santos, Ken Singleton, Pietro Veronesi, an anonymous referee, and seminar participants at Boston University, Duke Fuqua, MIT Sloan, Penn State Smeal, USC Marshall, the 2009 Adam Smith Asset Pricing Conference, the 2009 European Finance Association (Bergen), and the Winter 2010 Econometric Society Meetings (Atlanta) for comments. Tran Ngoc-Khanh and Yu Xu provided excellent research assistance. All the remaining errors are our own. Send correspondence to Scott Joslin, Assistant Professor of Finance and Business Economics, USC Marshall School of Business, Bridge Hall 308, 3670 Trousdale Parkway, Los Angeles, CA 90089; telephone: 213-740-7137. E-mail: sjoslin@usc.edu.

doi:10.1093/rfs/hhs065

and Markov-switching affine processes. In our primary result, we provide closed-form expression (up to an integral) for the following transform:

$$E_t^m\left[\exp\left(-\int_t^T R(X_s,s)ds\right)f(X_T)g(\beta\cdot X_T)\right], \tag{1}$$

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). Here, R is an arbitrary affine function that may represent the short rate or other types of discounting. We use E^m to denote the expectation under an arbitrary measure m. The substantial flexibility in the nonlinear expectation in Equation (1), as well as the process for X_t to allow time-varying conditional means, volatilities, and jump intensities, makes the above transform useful in dealing with generic nonlinearity problems in asset pricing (nonlinear stochastic discount factors or payoffs), estimation (nonlinear moments), and other areas.

Our method utilizes 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 allows us to replace any nonlinearities with an average of log-linear functions, for which the conditional characteristic function can be used to directly compute the expectations. This combination brings tractability to our generalized transform by avoiding intermediate Fourier inversions. Our method allows for a large class of nonlinear functions (tempered distributions), which include discontinuous and nondifferentiable functions as well as unbounded and nonintegrable functions, for which a standard Fourier transform might not exist. Moreover, a large variety of stochastic processes have tractable characteristic functions (such as affine jump-diffusions or Lévy processes). As a result, our method is applicable to a wide range of problems.

One area in which the generalized transform in Equation (1) can be a useful tool is risk-neutral pricing. We can value a large class of nonlinear payoffs analytically, provided that we know the forward conditional characteristic function of the underlying state variables.

The generalized transform can be also useful in economic modeling. For example, suppose that we want to value an asset under the historical measure ($\mathbb{P}$) using the stochastic discount factor m_t. The value at time t of a stochastic payoff y_T at time T is

$$P_t=\frac{1}{m(t,X_t)}E_t^{\mathbb{P}}[m(T,X_T)y(X_T)].$$

In the background, there is an equilibrium model that endogenously determines the stochastic discount factor m and payoff y as functions of the state variables X. In order to maintain tractability, we are often forced to adopt special utility functions (e.g., logarithmic utility), impose strong restrictions on the state variable process (e.g., i.i.d. or conditionally Gaussian), or log-linearize the model to obtain approximate solutions. The forward conditional characteristic function of X is typically quite complicated for general stochastic discount factors, which means that changing to the risk-neutral probability measure will not help simplify the problem.[1]

However, the above pricing equation bears resemblance to the generalized transform (1). The difficulties in pricing are typically due to the nonlinearity in the discount factor m or payoff y, which can be addressed using the new tools provided in this article. Through the generalized transform, we can (a) price assets with payoffs that are potentially discontinuous or nondifferentiable; (b) allow for more general preferences; (c) have more flexibility when introducing heterogeneity across agents, firms, or countries (in models of international finance); and (d) significantly enrich the underlying stochastic uncertainties governing the economy by introducing features such as time-varying growth rates, stochastic volatility, jumps, or cointegration restrictions.

To illustrate the application in risk-neutral pricing, we study the pricing of defaultable bonds with stochastic recovery rates. The recovery rate of a defaulted security can depend on firm characteristics and macroeconomic conditions. We provide analytic pricing for defaultable bonds under these conditions. We show that the negative correlation between recovery and default rates found in the data implies substantial nonlinearities in credit spreads. The comparison between the stochastic recovery model and widely used constant recovery models shows that ignoring stochastic recovery can lead to economically significant pricing errors.

In a second example, we study a general equilibrium model with time-varying labor income risk. We build upon the work of Santos and Veronesi (2006), who find that the share of labor income to consumption predicts future excess returns of the market portfolio. They attribute this result to the "composition effect": a higher labor share implies a lower covariance between consumption and dividends, which lowers the equity premium. Motivated by the empirical evidence—that volatilities of labor income and dividends as well as the correlation between the two change over time—we build a model with time-varying covariance between labor income and dividends. We obtain analytical solutions of the model via the generalized transform. In the calibrated model, we find that the composition effect is strong when the correlation between labor income and dividends is low (close to zero), but disappears when the correlation becomes significantly negative. In addition, the model has interesting implications for the comovement between the risk premium on financial wealth and human capital. Variations of this model can be used to study areas such as the cross-section of stock returns or international asset pricing.

[1] More precisely, the characteristic function under $\mathbb{Q}$ may be known to only an integral. Computing the expected nonlinear payoffs by risk-neutral pricing, the transform would then require a double integral, substantially increasing the computational difficulty.

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 particular, the transform analysis of general affine jump-diffusions in Duffie, Pan, and Singleton (2000, hereafter DPS) makes it easy to compute certain moments arising from asset pricing, estimation, and forecasting. Example applications include Singleton (2001), Pan (2002), Piazzesi (2005), and Joslin (2010), among others.[2]

When the moment functions do not conform to the basic DPS transform, one possible solution is to first recover the conditional density of the 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 by directly integrating over the density. Through this method, DPS obtain the extended transform for affine jump-diffusions, which they apply to option pricing.[3] Alternative methods to compute nonlinear moments include simulations or numerically solving the partial differential equations arising from the expectations via the Feynman-Kac methodology, which can be time-consuming and lacking accuracy, especially in high-dimensional cases. In our approach, we consider affine jump-diffusions and indeed any process with known conditional characteristic functions. Moreover, our method allows direct computation of a large class of nonlinear moments without the need to compute the (forward) density of the underlying state variable.

Bakshi and Madan (2000) connect the pricing of a class of derivative securities to the characteristic functions for a general family of Markov processes. In addition, they propose to approximate a nonlinear moment function with a polynomial basis, provided the function is entire, which in turn can be computed via the conditional characteristic function and its derivatives. Our method applies to more general nonlinear moment functions through the Fourier transform. We also extend the results to multivariate settings.

A few earlier studies have considered related Fourier methods. 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 (2011) takes the Fourier transform of a nonlinear pricing kernel that arises in the two-tree model of Cochrane, Longstaff, and Santa-Clara (2008).[4] In both studies, the state variables have i.i.d. increments.

[2] Gabaix (2009) considers a class of linearity generating processes, where particular moments (or accumulated moments) can have a very simple linear form given minor deviations from the assumption of affine dynamics of the state variable. This feature makes it very convenient to obtain simple formulas for the prices of stocks, bonds, and other assets. His work generalizes the model of Menzly, Santos, and Veronesi (2004).

[3] Other articles that take this approach include Heston (1993), Chen and Scott (1995), Bates (1996), Bakshi and Chen (1997), Bakshi, Cao, and Chen (1997), Chacko and Das (2002), Dumas, Kurshev, and Uppal (2009), and Buraschi, Trojani, and Vedolin (2010), among others.

[4] In the N-tree case, $N > 2$, Martin (2011) also provides an $(N-2)$-dimensional integral to compute the associated $(N-1)$-dimensional transform.

We view our approach as generalizing the related literature, in terms of both the moment function (to the class of tempered distributions) and the process of underlying state variables (including affine and Lévy processes).

1. Illustrative Example

Before presenting the main result of the article, we first illustrate the idea behind the generalized transform using an example of forecasting the average recovery rate of defaulted corporate bonds. The amount an investor recovers from a corporate bond upon default can depend on many factors, such as firm-specific variables (debt seniority, asset tangibility, and accounting information), industry variables (asset specificity and industry-level distress), and macroeconomic variables (aggregate default rates and business-cycle indicators). In addition, the recovery rate as a fraction of face value should in principle only take values from $[0,1]$.

A simple way to capture these features is to model the recovery rate using the logistic model

$$\varphi(X_t)=\frac{1}{1+e^{-\beta_0-\beta_1\cdot X_t}}, \tag{2}$$

where X_t is a vector of the relevant explanatory variables observable at time t. For example, Altman et al. (2005) model the aggregate recovery rate as a logistic function of the aggregate default rate, total amount of high-yield bonds outstanding, GDP growth, market return, and other covariates.

Investors may be interested in forecasting future average recovery rates in the economy. That is, we are interested in computing $E_0[\varphi(X_T)]$.[5] To simplify notation, we first define $Y_T\equiv\frac{1}{2}(-\beta_0-\beta_1\cdot X_T)$. We then rewrite

$$E_0[\varphi(X_T)]=E_0\left[\frac{1}{1+e^{2Y_T}}\right]=E_0\left[\frac{1}{2}e^{-Y_T}\frac{1}{\cosh(Y_T)}\right], \tag{3}$$

where we use the hyperbolic cosine function $\cosh(y)=\frac{1}{2}(e^y+e^{-y})$.

Although only a single variable Y_T appears in Equation (3), its conditional distribution may depend on the current values of each individual covariate, i.e., Y_t itself may not be Markov. Even if the covariates X_t follow a relatively simple process, direct evaluation of this expectation requires computing a multidimensional integral, which can be difficult when the number of covariates is large.

Suppose, however, that the conditional characteristic function (CCF) of X_T is known:

$$CCF(T,u;X_0)=E[e^{iu\cdot X_T}\,|\,X_0], \tag{4}$$

where $i=\sqrt{-1}$. If we could "approximate" the nonlinear term $1/\cosh(Y_T)$ inside the expectation of Equation (3) with exponential linear functions of

---

[5] We suppress the conditioning on the default event occurring at time T and further suppose that default occurring at time T is independent of the path $\{X_\tau\}_{0\le\tau\le T}$. This is stronger than the standard doubly stochastic assumption; relaxing this assumption is relatively straightforward but would complicate the example.

Y_T, then we would be able to use the characteristic function to compute the nonlinear expectation. As we elaborate in Section 2, this is achieved using the Fourier inversion of $1/\cosh(y)$,

$$\frac{1}{\cosh(y)} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \hat{g}(s) e^{isy} ds, \tag{5}$$

where $\hat{g}$ is the Fourier transform of $1/\cosh(y)$, which is known analytically (see, e.g., Abramowitz and Stegun 1964, 6.1.30 and 6.2.1), $\hat{g}(s) = \pi/\cosh(\frac{\pi s}{2})$.

Thus, we can substitute out $1/\cosh(Y_T)$ from Equation (3) and obtain

$$\begin{aligned} E_0[\varphi(X_T)] &= E_0\left[\frac{1}{2}e^{-Y_T}\frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{g}(s)e^{is\cdot Y_T}ds\right] \\ &= \frac{1}{4\pi}\int_{-\infty}^{+\infty}\frac{\pi}{\cosh(\frac{\pi s}{2})}E\left[e^{(-1+is)\cdot Y_T}\,\middle|\,X_0\right]ds \\ &= \frac{1}{4}\int_{-\infty}^{+\infty}\frac{1}{\cosh(\frac{\pi s}{2})}e^{-\frac{1}{2}(-1+is)\beta_0}CCF\left(T,-\frac{1}{2}(i+s)\beta_1;X_0\right)ds, \end{aligned} \tag{6}$$

where for the second equality we assume that the order of integrals can be exchanged; the third equality follows from applying the result in Equation (4). Now all that remains for computing the expected recovery rate is to evaluate a one-dimensional integral (regardless of the dimension of X_t), which is a significant simplification compared to the direct approach.

If X_t follows an affine process, its conditional characteristic function takes a particularly simple form; it is an exponential affine function of X_0. In some cases, the exact form is known in closed form. Even in the general case in which no closed-form solution is available, the affine coefficients are simple to compute as the solutions to differential equations. In those cases, our approach again offers a great deal of simplicity in the face of a possible curse of dimensionality; the linear scaling involved in solving N ordinary differential equations is dramatically easier than solving an N-dimensional partial differential equation.

2. Generalized Transforms

We now present our theoretical results. Our results only require that the conditional characteristic functions of the underlying state variables are tractable, and our results apply whether the stochastic process is modeled in discrete or continuous time. We will start with the case of continuous time affine jump-diffusion (AJD), because its conditional characteristic function is particularly easy to compute and AJDs have been widely used in economics and finance. In Section 2.2, we discuss examples of processes that are not continuous-time affine jump-diffusions.

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 2004). Suppose that X is a Markov process in some state space $D \subset \mathbb{R}^N$ satisfying the stochastic differential equation

$$dX_t = (K_0 + K_1 X_t)dt + \sigma_t dW_t + dZ_t, \tag{7}$$

where W is an $\mathcal{F}_t$-standard n-dimensional Brownian motion, Z is a pure jump process with arrival intensity $\lambda_t = \ell_0 + \ell_1 \cdot X_t$ and fixed D-invariant distribution ν, and $(\sigma_t \sigma_t')_{i,j} = H_{0,ij} + \sum_k H_{1,ijk} X_t^k$ with $H_0 \in \mathbb{R}^{N\times N}$ and $H_1 \in \mathbb{R}^{N\times N\times N}$. 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 $(K_0, K_1, H_0, H_1, \ell_0, \ell_1, \nu, \rho_0, \rho_1)$.

To establish our main result, let us first review some basic concepts from distribution theory. A function $f : \mathbb{R}^N \to \mathbb{R}$, which is smooth and rapidly decreasing in the sense that for any multi-index α and any $P \in \mathbb{N}$, $\|f\|_{P,\alpha} \equiv \sup_x |\partial^\alpha f(x)|(1+\|x\|)^P < \infty$, is referred to as a Schwartz function. Here, $\partial^\alpha f$ denotes the higher-order mixed partial of f associated with the multi-index $\alpha = (\alpha_1, \ldots, \alpha_N)$ (i.e., $\partial^\alpha f = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_N}}{\partial x_N^{\alpha_N}} f$), and $\|x\|$ is the Euclidean norm of the vector x. The collection of all Schwartz functions is denoted $\mathcal{S}$, and $\mathcal{S}$ is endowed with the topology generated by the family of semi-norms $\|f\|_{P,\alpha}$. 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 g that 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 $f \mapsto \langle g, f \rangle$, where we use the inner-product notation

$$\langle g, f \rangle = \int_{\mathbb{R}^N} g(x) f(x) dx. \tag{8}$$

As is standard, we maintain the inner product notation, even when a tempered distribution g does not correspond to a function as in Equation (8). For example, the δ-function is a tempered distribution given by $\langle \delta, f \rangle = f(0)$, which does not arrive from a function. Throughout the article, we use the notation $\delta(s)$ to denote the Dirac delta function, with $\delta_x(s) = \delta(s - x)$.

For our considerations, the key property is that the set of tempered distributions is suitable for Fourier analysis. For any Schwartz function f, the Fourier transform of f is another Schwartz function, denoted $\hat{f}$, and is defined by

$$\hat{f}(s) = \int_{\mathbb{R}^N} e^{-is\cdot x} f(x) dx. \tag{9}$$

The Fourier transform can be inverted through the relation

$$f(x) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} e^{ix\cdot s} \hat{f}(s) ds, \tag{10}$$

which holds pointwise for any Schwartz function. The Fourier transform is extended to apply to tempered distributions through the definition $\langle \hat{g}, f \rangle =$

$\langle g, \hat{f} \rangle$. This extension is useful because many functions define tempered distributions (through Equation [8]) but do not have Fourier transform in the sense of Equation (9) because the integral is not well defined. An example is the Heaviside function:

$$H(x) = 1_{\{0 \le x\}} \Rightarrow \hat{H}(s) = \pi \delta(s) - \frac{i}{s}, \tag{11}$$

where integrating against $1/s$ is to be interpreted as the principal value of the integral. Considering distributions allows us to consider functions that 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 $g \in \mathcal{S}^*$ and (Θ, α, β) satisfies Assumptions 1 and 2 in Appendix 1. Then,

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

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

$$s \mapsto \psi(\alpha + s\beta i) = E_0\left[e^{-\int_0^T R(X_u)du} e^{(\alpha + is\beta) \cdot X_T}\right]. \tag{13}$$

In the case in which X follows an affine jump-diffusion, the discounted conditional characteristic function ψ is given in DPS,

$$\psi(\alpha + is\beta) = e^{A(T; \alpha + is\beta, \Theta) + B(T; \alpha + is\beta, \Theta) \cdot X_0}, \tag{14}$$

and A, B are solutions to a system of ordinary differential equations (ODEs) that can generally be computed easily (see Appendix 1 for more details).

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

$$H = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{g}(s) \psi(\alpha + is\beta) ds. \tag{15}$$

Bakshi and Madan (2000) show that option-like payoffs (affine translations of integrable functions) can be spanned by a continuum of characteristic functions. Theorem 1 above shows that the characteristic functions span a much larger set of functions, which includes all tempered distributions. Equation (15) makes this spanning explicit for the cases in which $\hat{g}$ defines a function.

There is some flexibility in the choice of α and g in Equation (12). Notice that

$$e^{\alpha \cdot X_T} g(\beta \cdot X_T) = e^{(\alpha - c\beta) \cdot X_T} \tilde{g}(\beta \cdot X_T), \tag{16}$$

where $\tilde{g}(s) = e^{cs} g(s)$. This property can be useful in the case in which g is not integrable but decreases rapidly as s approaches either positive or negative infinity (e.g., the logit function). In this case, such a transformation of g makes it possible to apply (12).

2.1 Two extensions

The result of Theorem 1 can be extended in a number of ways. First, we introduce a class of *pl-linear* (polynomial-log-linear) functions:

$$f(\alpha,\gamma,p,X)=\sum_i p_i X^{\gamma_i} e^{\alpha_i\cdot X}, \tag{17}$$

where $\{p_i\}$ are arbitrary constants, $\{\alpha_i\}$ are complex vectors, and $\{\gamma_i\}$ are arbitrary multi-indices so that $X^\gamma=\prod_j X_j^{\gamma_i}$. For example, with $N=3$ and $\gamma=(1,2,1)$, $X^\gamma=X_1^1X_2^2X_3^1$. The following proposition extends Theorem 1 to work with any *pl-linear* functions.

Proposition 1. Suppose that $g\in\mathcal{S}^*$, and $(\Theta,\alpha,\beta,\gamma)$ satisfies Assumption 1' and Assumption 2' in Appendix 2. Then,

$$\begin{aligned}H(g,\alpha,\beta,\gamma)&=E_0\left[\exp\left(-\int_0^T R(X_u)du\right)X_T^\gamma e^{\alpha\cdot X_T}g(\beta\cdot X_T)\right]\\&=\frac{1}{2\pi}\langle\hat{g},\psi(\alpha+\cdot\beta i;\gamma)\rangle,\end{aligned} \tag{18}$$

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

$$s\mapsto\psi(\alpha+s\beta i;\gamma)=E_0\left[e^{-\int_0^T R(X_u)du}X_T^\gamma e^{(\alpha+is\beta)\cdot X_T}\right]. \tag{19}$$

The function ψ is computed by solving the associated ODE in Appendix 2.

It is immediate from 1 that we can now compute expectations of the form

$$H(f,g,\alpha,\beta)=E_0\left[\exp\left(-\int_0^T R(X_u)du\right)f(\alpha,\gamma,p,X_T)g(\beta\cdot X_T)\right]. \tag{20}$$

The assumption that the function g in the generalized transform is 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 or 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 $g(\beta_1\cdot X,\cdots,\beta_M\cdot X)$ for $M\in\mathbb{N}$.

Proposition 2. Suppose that $g\in\mathcal{S}_M^*$ (an M-dimensional tempered distribution), $\alpha\in\mathbb{R}^N$, $\mathfrak{b}\in\mathbb{R}^{M\times N}$, and $(\Theta,\alpha,\mathfrak{b})$ satisfies Assumptions 1 and 2 in

Appendix 1. Then,

$$H(g,\alpha,\mathfrak{b})=E_0\left[\exp\left(-\int_0^T R(X_s)ds\right)e^{\alpha\cdot X_T}g(\mathfrak{b}X_T)\right]$$
$$=\frac{1}{(2\pi)^M}\langle\hat{g},\psi_M(\alpha+\cdot\mathfrak{b}i)\rangle, \tag{21}$$

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

$$\psi_M:\mathbb{C}^M\to\mathbb{C}\ ,\ s\mapsto\psi_M(\alpha+s^\top\mathfrak{b}i)=E_0\left[e^{-\int_0^T r(X_u)du}e^{(\alpha+is^\top\mathfrak{b})\cdot X_T}\right]. \tag{22}$$

It is immediate to extend the transform in Proposition 2 by replacing $e^{\alpha\cdot X_T}$ with a *pl-linear* function as in Proposition 1.

Fourier transforms of many functions are known in closed form (see, e.g., Folland 1984). Additionally, standard rules allow for differentiation, integration, product, convolution, and other operations to be conducted while maintaining closed-form expressions. Even if the function $\hat{g}$ is not known in closed form, including those cases in which g itself is given as an implicit function, it is straightforward to compute numerically (a one-dimensional integral in the case of Theorem 1 and Proposition 1, an M-dimensional integral in Proposition 2). Alternatively, one might consider approximating g with a function $\widetilde{g}$ for which the Fourier transform is known in closed form.

For a given set of parameters, the Fourier transform of g and the coefficients A and B in Equation (14) need only be computed once. Once computed, the Fourier transform and the differential equation solutions can be used repeatedly to compute moments with different initial values of the state variable X_0 or horizon T. When the moment function takes the form of $f(\alpha,\gamma,p,X)g(\beta\cdot X)$ as in Proposition 1, the same Fourier transforms and differential equation solutions can also be used to compute moments with different *pl-linear* function f.

2.2 Beyond affine jump-diffusions

The key aspects of Theorem 1 and the two extensions are the ability to compute the transform given in Equations (13), (19), or (22). These transforms are tractable for affine jump-diffusions. However, other stochastic processes can also be suitable for the generalized transform, provided that the appropriate (forward) conditional characteristic function can be computed. One example is the discrete-time affine processes. Appendix 3 presents the generalized transform result in discrete time.

Another example is the class of Lévy processes (see, e.g., Protter 2004). Lévy processes allow for both finite and infinite activity jumps, though in some contexts the assumption of independent increments may be restrictive. To be concrete, consider, e.g., the process of Carr et al. (2002). They specify a pure jump Lévy process with Lévy measure ν given by the density k_{CGMY}, where

$$k_{CGMY}(x)=\begin{cases}Ce^{-G|x|}|x|^{-1-Y} & \text{if } x>0\\ Ce^{-M|x|}|x|^{-1-Y} & \text{otherwise,}\end{cases}$$

where (C,G,M,Y) are constants. When $Y=-1$, this reduces to i.i.d. jump arrivals with an exponential distribution. When $Y=0$, we recover the variance gamma process studied by Madan, Carr, and Chang (1998). Generally, the CGMY process allows for flexibility in modeling the activity of small and large jumps, as well as in the tail properties for large jumps.

The Lévy-Khintchine formula allows us to recover the conditional characteristic function from an arbitrary Lévy measure. Carr et al. (2002) show that if X_t is CGMY process, then the characteristic function is

$$E_0[e^{iu(X_t-X_0)}]=\exp\big(tC\Gamma(-Y)[(M-iu)^Y-M^Y+(G+iu)^Y-G^Y]\big), \quad (23)$$

where Γ denotes the standard Gamma function $\Gamma(t)=\int_0^\infty s^{t-1}e^{-s}ds$. Using the above characteristic function, we can then apply the results of Theorem 1 and Proposition 1 to compute nonlinear moments of Lévy processes.

Other examples of non-AJD processes with tractable conditional characteristic functions include the Markov-switching affine process and the discrete-time autoregressive process with gamma-distributed shocks. As shown in Dai, Singleton, and Yang (2007) and Ang, Bekaert, and Wei (2008), one can incorporate regime shifts in the conditional mean, conditional covariance, or conditional probability of jumps into standard AJDs. Bekaert and Engstrom (2010) study a general equilibrium model that uses autoregressive processes with gamma-distributed shocks. In both examples, the conditional characteristic functions can be computed easily.

Having presented the theory of the generalized transform, we next illustrate its power in pricing contingent claims and solving equilibrium asset pricing models.

3. Applications in Contingent Claim Pricing

An example we study in this section is pricing defaultable bonds with stochastic recovery rates. Recovery rates may depend nonlinearly on the state of the economy, which potentially makes the recovery function nonintegrable or nonsmooth. Our method can easily handle such models. We show that stochastic variations in the recovery rates can not only have large effects on the average level of credit spreads but also lead to economically important nonlinearities in credit spreads as default intensity changes, which are difficult to capture using standard models with constant recovery rate. We complete the section by comparing our method for risk-neutral pricing of contingent claims with some alternative methods.

3.1 Stochastic recovery

Following up on the illustrating example in Section 1, we now examine the pricing of credit-risky securities (e.g., defaultable bonds or credit default swaps) with stochastic recovery upon default. The importance of stochastic recovery in credit risk modeling is supported by the empirical evidence of

Altman et al. (2005), Acharya, Bharath, and Srinivasan (2007), and Chen (2010), among others, who show that the recovery rates of corporate bonds are significantly related to industry distress and macroeconomic conditions. Our main goal for this exercise is to answer the following question: Does the nonlinearity in corporate bond prices, introduced by stochastic recovery, matter economically?

Consider a T-year defaultable zero-coupon bond with its face value normalized to 1. Following the literature of reduced-form credit risk models, the default time is assumed to be a stopping time τ with risk-neutral intensity λ_t.[6] The risk-neutral recovery rate at default $\widetilde{\varphi}_t$ is a bounded predictable process that is adapted to the filtration $\{\mathcal{F}_t : t \geq 0\}$. The instantaneous risk-free rate is r_t. Then, the price of the bond is

$$\begin{aligned} V_t &= E_t^{\mathbb{Q}}\left[e^{-\int_t^\tau r_u du} 1_{\{\tau \leq T\}} \widetilde{\varphi}_\tau\right] + E_t^{\mathbb{Q}}\left[e^{-\int_t^T r_u du} 1_{\{\tau > T\}}\right] \\ &= E_t^{\mathbb{Q}}\left[\int_t^T e^{-\int_t^s (r_u+\lambda_u)du} \lambda_s \widetilde{\varphi}_s ds + e^{-\int_t^T (r_u+\lambda_u)du}\right]. \end{aligned} \tag{24}$$

The second equality follows from the doubly stochastic assumption and regularity conditions. Let X_t be the vector of state variables that determine the risk-free rate, default intensity, and recovery rate. Suppose that both r_t and λ_t are affine in X_t. In addition, we model the risk-neutral recovery rate as $\widetilde{\varphi}_t = g(\beta \cdot X_t)$ for some proper function g, which ensures that the recovery rate is between 0 and 1 in addition to satisfying a suitable no-jump condition.[7] Importantly, the risk-neutral recovery rate may differ from the historical recovery rate when investors demand a premium for bearing recovery risk.

To investigate the quantitative impact of stochastic recovery on the pricing of defaultable bonds, we directly specify the dynamics of state variables $X_t = [\lambda_t \ Y_t]'$ under the risk neutral measure $\mathbb{Q}$ as follows:

$$d\lambda_t = \kappa_\lambda(\theta_\lambda - \lambda_t)dt + \sigma_\lambda \sqrt{\lambda_t} dW_t^\lambda, \tag{25}$$

$$dY_t = \kappa_Y(\theta_Y - Y_t)dt + \sigma_Y \sqrt{\lambda_t} dW_t^Y, \tag{26}$$

where W_t^λ and W_t^Y are uncorrelated Brownian motions. The risk-free rate is given by

$$r_t = Y_t - \delta\lambda_t. \tag{27}$$

This simple setup (with $\delta > 0$) captures the negative correlation between r_t and λ_t in the data.

6 To be precise, we fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and two filtrations $\{\mathcal{F}_t : t \geq 0\}$, $\{\mathcal{G}_t : t \geq 0\}$. The default time is a totally inaccessible $\mathcal{G}$-stopping time $\tau : \Omega \to (0, +\infty]$. We assume that under the risk-neutral measure $\mathbb{Q}$, τ is doubly stochastic driven by the filtration $\{\mathcal{F}_t : t \geq 0\}$. See Duffie (2005) for a survey on the reduced form approach for modeling credit risk and the doubly stochastic property.

7 The 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.

With the help of the generalized transform, we now have flexibility in choosing the recovery function $\widetilde{\varphi}_t$ and still maintain tractability for pricing. For example, the cumulative distribution function (CDF) of any probability distribution, such as the logit or probit models, will take values in $[0,1]$ and can be used to model the recovery rate. Modeling $\widetilde{\varphi}$ with CDFs has the added benefit of having nice Fourier transform properties. For example, the integrands of the Gaussian and Cauchy models have a closed-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 the CDF in such cases.[8]

For simplicity, we assume that $\widetilde{\varphi}$ only depends on the default intensity, and we adopt a modified Cauchy model:

$$\widetilde{\varphi}(\lambda)=\frac{a}{1+b(\lambda-\lambda_0)^2}+c. \tag{28}$$

The constant term $c\in[0,1]$ sets a lower bound for $\widetilde{\varphi}$ that is potentially above 0, which gives us more flexibility in matching the empirical distribution of recovery rates. The Fourier transform of $\widetilde{\varphi}$ (excluding the constant c) is

$$\hat{\widetilde{\varphi}}(t)=\frac{a\pi}{\sqrt{b}}e^{\lambda_0 it-\frac{1}{\sqrt{b}}|t|}. \tag{29}$$

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

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

which is mapped into the generalized transform of Theorem 1 by choosing

$$f(\alpha\cdot X)=\iota_1\cdot X,$$

$$g(\beta\cdot X)=\frac{a}{1+b(\iota_1\cdot X-\lambda_0)^2}+c,$$

where $\iota_1=[1\ 0]'$. Notice that one can introduce additional state variables in X to capture richer dynamics of the term structure, default risk, recovery rate, and macroeconomic conditions, which does not add any complication to pricing as long as the risk-neutral recovery rate is still given by $\widetilde{\varphi}_t=g(\beta\cdot X_t)$.

We now use the processes of default intensity λ_t and risk-free rate r_t (25)–(27) and the recovery model (28) to price a five-year defaultable zero-coupon bond. We calibrate 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.

[8] There are also specifications for which the existing methods apply, e.g., $\widetilde{\varphi}(X)=e^{\beta\cdot X}1_{\{\beta\cdot X<0\}}+1_{\{\beta\cdot X>0\}}$. Bakshi, Madan, and Zhang (2006) study such a setting. However, our method is more general. For example, a power law specification, such as $\widetilde{\varphi}(X)=(\beta\cdot X)^{-\alpha}+c$ for $\alpha,c>0$, falls under our theory but would not be directly solvable with existing methods.

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

We consider two calibrations of $\widetilde{\varphi}(\lambda)$ in Equation (28). First, we directly calibrate the risk-neutral recovery rate to the actual recovery rates. Using Moody's data on aggregate annual default rates and recovery rates, we calibrate $a=0.68$, $b=2000$, $c=0.25$, and $\lambda_0=-0.014$. The fitted function is Model 1 in Figure 1. Fitting the risk-neutral recovery rate to the historical recovery rate amounts to assuming no recovery risk premium, and by treating the physical and risk-neutral default intensity as the same, we are also assuming that there is no jump-to-default risk premium. To assess the impact that a recovery risk premium would have on corporate bond spreads, we also consider a second calibration, which incorporates a risk adjustment to the historical recoveries. The fitted function is Model 2 in Figure 1, which has very similar recovery rates to Model 1 when the default intensity is low but has a sharper decline in recovery rates than does Model 1 as default intensity rises. The widening gap between the two models implies that the recovery risk premium in Model 2 is increasing with the aggregate default probability. In the second calibration, we assume $a=0.9$, $b=1200$, $c=0$, and $\lambda_0=-0.014$.[9]

A standard assumption for default recovery in both academic analysis and industry practice is that the risk-neutral recovery rate is constant, which is often set to 25% (see, e.g., Pan and Singleton 2008). This value is lower than the historical mean recovery rate, which is a parsimonious way to capture the recovery risk premium. We compare our stochastic recovery model with a model with 25% constant recovery rate. In addition, we consider another constant recovery rate, which is calibrated to best fit the stochastic recovery models (by minimizing the mean square error computed based on the empirical distribution).[10]

In Figure 2, the top panels investigate the Stochastic Recovery Model 1 (no recovery risk premium), whereas the bottom panels consider the Stochastic Recovery Model 2 (risk adjustment for recovery risk). All the results are computed with the risk-free rate fixed at 5%. We see from Panels A and C that the credit curve is almost linear in the default intensity for models with constant recovery rates.[11] Changing the recovery rate from 25% to another value

[9] In principle, one could identify the magnitude of the recovery premium empirically; here, we simply want to assess the order of magnitude of the differences in corporate credit spreads from a (linear) model with constant recoveries and a (nonlinear) model with stochastic recoveries.

[10] We have also considered the analogous recovery of market value models, but as in Duffie and Singleton (1999), we found that in our calibration the credit spreads in the constant recovery of market value models closely match those in the constant recovery of face value models, so we omit them from our comparison.

[11] For a recovery of market value model, the yields are exactly affine in the default rate, as shown by Duffie and Singleton (1999).

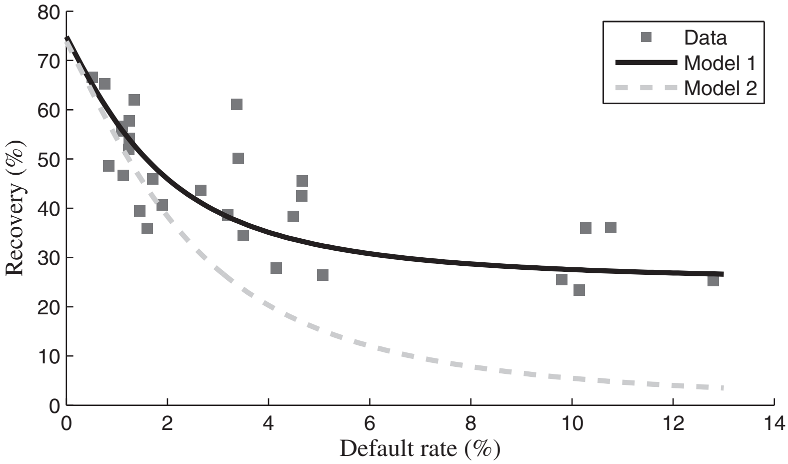

Figure 1
A Cauchy model of risk-neutral recovery rates
This figure plots the historical aggregate recovery rates and default rates from Altman and Kuehne (2011) for the period 1982–2010. The solid line is a Cauchy recovery model fitted to the historical data, which assumes that there is no recovery risk premium. The dashed line is a Cauchy recovery model with recovery risk premium so that the risk-neutral expected recovery rates are below the historical rates in aggregate bad states of the economy.

primarily results in a change in the steepness of the credit curve with respect to the default intensity. In contrast, the stochastic recovery curves exhibit a fair amount of nonlinearity, with some convexity in the region of small default probabilities. The convexity is caused by the fact that when default intensities are low (high), recoveries are high (low), so the incremental effect of an increase in default probabilities is small (large); thus, the curve becomes steeper as the intensity increases. This nonlinearity in the credit curve (as a function of default intensity) is a qualitative difference between the stochastic and constant recovery models.

Panels B and D assess the economic importance of stochastic recovery by plotting the pricing errors of constant recovery models relative to the stochastic recovery models. We see that there are quite large differences between the fixed recovery rate of the standard 25% recovery specification and the two stochastic recovery models; the root-mean-square differences across the given range of intensities are 38 and 206 bps for Models 1 and 2, respectively. On the one hand, relative to Model 1 (without recovery risk premium), the 25% constant recovery assumption can lead to overstating the credit spread by up to 50 bps for moderate default intensities. On the other hand, relative to Model 2, a 25% constant recovery rate becomes too conservative and leads us to understate the credit spread most of the time, where the pricing errors can be as large as 400 bps.

Recalibrating the constant recovery rate for each model produces a somewhat closer fit. The optimized recovery rates are 29.6% for Model 1 and 8.5% for Model 2, respectively, which result in root-mean-square differences of 26 and

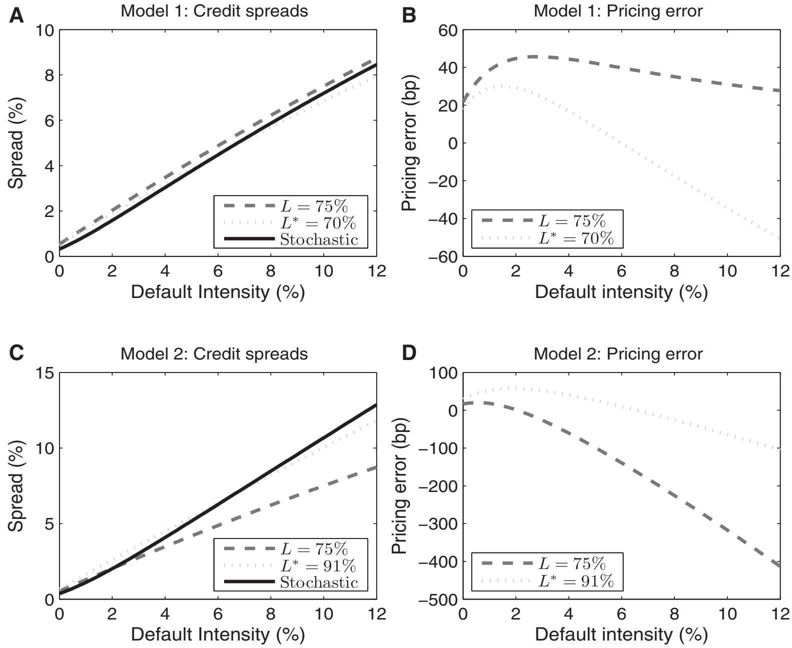

Figure 2
Credit spreads for five-year bonds with constant recovery and Cauchy recovery
For different values of conditional default intensity, Panels A and C plot the credit spreads of a 5-year zero-coupon defaultable bond as implied by Stochastic Recovery Models 1 and 2, as well as the spreads implied by two constant recovery models (one with default loss $L=75\%$, the other that provides the best fit to the corresponding stochastic recovery model). Panels B and D plot the pricing errors of the two constant recovery models relative to the stochastic recovery models.

51 bps relative to Models 1 and 2. However, the yield differences are still economically significant most of the time, with the constant recovery model typically overstating the credit spread for low default intensity and understating the spread for high intensity.

In summary, our analysis shows that quantitatively, the nonlinear effect induced by stochastic recovery is large relative to the standard model with constant recovery. It demonstrates that (1) within the class of constant recovery models, the industry standard 25% recovery rate is often substantially different from the optimal recovery rate; and (2) even the ex post optimal recovery rate produces sizable pricing errors relative to the stochastic recovery model, especially when the recovery risk premium is high. These results highlight the importance of carefully incorporating nonlinearities caused by stochastic recovery rates into credit risk modeling.

3.2 Comparison with some alternative methods

We now discuss how the generalized transform method in this article differs from the methods developed by Duffie, Pan, and Singleton (2000) and Bakshi and Madan (2000). We will use the example of pricing an European option to illustrate some of the key differences.

Again, we assume that X_t is the vector of state variables that follows an affine process under the risk-neutral probability measure $\mathbb{Q}$ and that the instantaneous risk-free rate satisfies $r_t = \rho_0 + \rho_1 \cdot X_t$. Consider the example of an European put option with strike K. Let $\beta \cdot X_t$ be the log stock price, then the payoff is $(K - e^{\beta \cdot X_T})^+$. Define $g(\beta \cdot X) \equiv 1_{\{\beta \cdot X \le \log K\}}$. The option price can be written as

$$P_t = E_t^{\mathbb{Q}}\left[e^{-\int_t^T r_s ds + y} g(\beta \cdot X_T)\right] - E_t^{\mathbb{Q}}\left[e^{-\int_t^T r_s ds + \beta \cdot X_T} g(\beta \cdot X_T)\right],$$

which can be computed by applying Theorem 1. In this case, the Fourier transform of g is defined as a distribution:

$$\hat{g}_y(s) = \pi \delta(s) + \frac{i e^{-isy}}{s}, \tag{30}$$

where the second term is interpreted as a principal value integral. It follows that

$$\begin{aligned} E_t^{\mathbb{Q}}\left[e^{-\int_t^T r_s ds + \alpha \cdot X_T} g_y(\beta \cdot X_T)\right] &= \int_{-\infty}^{\infty} \left(\frac{\delta(s)}{2} - \frac{e^{-isy}}{2\pi i s} \right) \psi(\alpha + is\beta) ds \\ &= \frac{\psi(\alpha)}{2} - \int_0^{\infty} \text{Real}\left(\frac{\psi(\alpha + is\beta) e^{-isy}}{\pi i s} \right) ds. \end{aligned} \tag{31}$$

In the last equation, we use the fact that the real part of the integrand is even and the imaginary part is odd.

The above result replicates the formula given in DPS obtained by Lévy inversion. However, their results are limited to the case of affine jump-diffusion and payoffs that are *pl-linear* in the underlying state variable, which is a special case of Theorem 1. DPS arrive at this equation by effectively computing the forward density by Fourier transform (a one-dimensional integral) and then integrating over the payoff region (now a two-dimensional integral). In this case, Fubini and limiting arguments allow this two-dimensional integral to be reduced to a one-dimensional integral as in the standard Lévy inversion formula (without a forward measure).

Bakshi and Madan (2000, henceforth BM) provide a more general method for pricing options. It allows for payoffs of the form $(H(X_T) - K)^+$, where X is a univariate stochastic process with known conditional characteristic function under the risk-neutral measure, H is a positive and entire function (analytic at all finite points on the complex plane), and K is a fixed strike price. BM propose a power series expansion of H, the expectations of which can then be computed through differentiation of the conditional characteristic functions. Their method applies to nonaffine processes as well.

Our method differs from BM in several aspects. First, the power series expansion approach in BM requires that H be infinitely differentiable and that the power series converge to $H(X)$ for all X. However, there are cases in which the payoff function has many kinks or is not entire (simple examples include $\ln(X)$ and $\sqrt{X}$). Second, in those cases in which the Fourier transform $\hat{g}$ is known, our method requires only a one-dimensional integration, whereas the method of BM requires a one-dimensional integration and an infinite sum. Third, we extend BM's result on spanning of option payoff with characteristic functions. In BM, the set of payoff functions H that can be spanned by characteristic functions is an affine translation of an $\mathcal{L}^1$ function (see BM Theorem 1).[12] We relax the growth condition on the underlying payoff; the dual space $\mathcal{S}^*$ is quite large and contains $\mathcal{L}^p$ for any p, as well as functions not in $\mathcal{L}^p$ for any p. Finally, our theory extends the analysis to multivariate settings.

To illustrate some of these differences, consider the payoff of the form $H(X)=X^\alpha$ for a positive noninteger α. In this case, H is not entire (for any choice of the center of the power series, X_0, the power series converges to $H(X)$ only for $0<X<2X_0$). It is also not an affine translation of an $\mathcal{L}^1$ function. In such cases, if one were to compute option prices through the Taylor expansion of the nonentire payoff function with a particular choice of X_0 and the order of Taylor expansion, the error could be very large. However, we can still use the fact that H represents a tempered distribution to write

$$\hat{g}(s)=(is)^{-1-\alpha}\Gamma_{\text{inc}}(\alpha+1,isk)-Ke^{-iks}\hat{H}(s), \tag{32}$$

where $k=K^{\frac{1}{\alpha}}$, $\hat{H}$ is given by (11), $\Gamma_{\text{inc}}(\cdot,\cdot)$ denotes the incomplete Gamma function, and $s^{-1-\alpha}$ is interpreted in the sense of a homogeneous tempered distribution.

4. Applications in Economic Modeling

In addition to risk-neutral pricing, the generalized transform can also be employed in economic modeling. For example, the need to compute expectations of nonlinear functions of state variables arises naturally when we price assets using the stochastic discount factor under the physical measure $\mathbb{P}$. Suppose that the state variables are affine under $\mathbb{P}$ and the underlying economic model gives rise to a stochastic discount factor $m(t,X)$. Then, the present value of a stochastic payoff $Y_T=y(X_T)$ at time T is

$$P_t=\frac{1}{m(t,X_t)}E_t^{\mathbb{P}}[m(T,X_T)y(X_T)]. \tag{33}$$

Except for some special cases (when m_t is an exponential affine function of X_t), the dynamics of X_t under the risk-neutral measure $\mathbb{Q}$ can be quite complicated, and the risk-free rate r_t may not be affine in X_t, making it difficult to do

[12] A function H is of class $\mathcal{L}^p$ if $\left(\int|H(x)|^p dx\right)^{1/p}<\infty$.

pricing under $\mathbb{Q}$. Instead, prices can be more easily computed under $\mathbb{P}$ using the generalized transform, provided that $m(T,X)y(X)$ can be decomposed into $f(X)g(\beta\cdot X)$ as in Section 2.

Two classes of widely used models, where such non-*pl-linear* stochastic discount factors arise, are in general equilibrium models where there are heterogeneities (1) in the sources of income or in the cross-section of stocks or (2) across agents in terms of preferences or beliefs. We focus on the first case in this section and leave the analysis of models with heterogeneous agents to the online appendix (Chen and Joslin 2011) applications with differences in beliefs and disaster risk are in Chen, Joslin, Tran (2010) and Chen, Joslin, and Tran (2012).

Several recent articles have studied the general equilibrium effects of multiple sources of income (consumption) and demonstrated their importance for understanding the time series and cross-section of asset prices. See Santos and Veronesi (2006, analytical result for a model with financial asset and labor income that satisfy special cointegration restrictions), Piazzesi, Schneider, and Tuzel (2007, numerical solution for a model with housing and nonhousing consumption), Cochrane, Longstaff, and Santa-Clara (2008, analytical result for a model with two i.i.d. trees and log utility), and Martin (2011, analytical result for N i.i.d. trees and power utility). In the following example, we use the generalized transform to obtain analytical results in a model with power utility and non-i.i.d. trees.

Suppose that there are two assets (or two types of consumption goods) in the economy, both in unit supply, with dividends paying out continuously at rates $D_{1,t}$ and $D_{2,t}$. We assume that the log dividends $d_{1,t}=\log D_{1,t}$ and $d_{2,t}=\log D_{2,t}$ are part of a vector X_t ($d_{1,t}=\iota_1\cdot X_t, d_{2,t}=\iota_2\cdot X_t$, with $\iota_1=[1\ 0\ 0\cdots]'$ and $\iota_2=[0\ 1\ 0\cdots]'$), which follows an affine jump-diffusion (7). This model can allow for time variation in the expected dividend growth rates, stochastic volatility, and time variation in the probabilities of jumps. Cointegration restriction also can be imposed to allow for stationary of the shares of the two assets.

There is an infinitely-lived representative investor with constant relative risk aversion (CRRA utility) over aggregate consumption:

$$U(c)=E_0^{\mathbb{P}}\left[\int_0^\infty e^{-\rho t}\frac{C_t^{1-\gamma}-1}{1-\gamma}dt\right], \tag{34}$$

where γ is the coefficient of relative risk aversion and ρ is the time discount rate. Aggregate consumption is a CES aggregator of the two goods $D_{1,t}$ and $D_{2,t}$,

$$C_t=\left(D_{1,t}^{(\epsilon-1)/\epsilon}+\omega D_{2,t}^{(\epsilon-1)/\epsilon}\right)^{\epsilon/(\epsilon-1)}. \tag{35}$$

The parameter ϵ is the elasticity of intratemporal substitution between the two goods, and ω determines the relative importance of the two goods.

We can recover the continuous-time version of Piazzesi, Schneider, and Tuzel (2007) when we interpret $D_{1,t}$ and $D_{2,t}$ as housing and nonhousing consumption, respectively, and assume that the growth rate of nonhousing consumption is i.i.d. and the log ratio of the two dividends follows a square-root process. In the case $\epsilon \to \infty$ and $\omega = 1$, the two goods become perfect substitutes so that $C_t = D_{1,t} + D_{2,t}$. This is the case considered by Cochrane, Longstaff, and Santa-Clara (2008) and Martin (2011), both of which assume i.i.d. dividend growth.

In equilibrium, there is a unique stochastic discount factor $m_t = e^{-\rho t} C_t^{-\gamma}$. Under the standard regularity conditions, the price of asset i $(i = 1,2)$, $P_{i,t}$, is then given by

$$P_{i,t} = E_t^{\mathbb{P}}\left[\int_0^\infty \frac{m_{t+u}}{m_t} D_{i,t+u} du\right]$$
$$= \left(D_{1,t}^{(\epsilon-1)/\epsilon} + \omega D_{2,t}^{(\epsilon-1)/\epsilon}\right)^{\gamma\epsilon/(\epsilon-1)}$$
$$\int_0^\infty e^{-\rho u} E_t^{\mathbb{P}}\left[\frac{D_{i,t+u}}{\left(D_{1,t+u}^{(\epsilon-1)/\epsilon} + \omega D_{2,t+u}^{(\epsilon-1)/\epsilon}\right)^{\gamma\epsilon/(\epsilon-1)}}\right] du. \qquad (36)$$

The main challenge of computing the stock price comes from the stochastic discount factor, which is non-*pl-linear* in the state variable X_t. As a result, the risk-free rate is not affine in X_t, and X_t is not affine under the risk-neutral measure. To map the expectation in Equation (36) into the generalized transform, we rewrite the expectation inside the integral of (36) as

$$E_t^{\mathbb{P}}\left[\frac{D_{1,s}}{\left(D_{1,s}^{(\epsilon-1)/\epsilon} + \omega D_{2,s}^{(\epsilon-1)/\epsilon}\right)^{\gamma\epsilon/(\epsilon-1)}}\right] = E_t^{\mathbb{P}}\left[\frac{e^{(1-\gamma/2)d_{1,s} - \gamma/2 d_{2,s}}}{\left(2\cosh\left(\frac{\epsilon-1}{\epsilon}\frac{d_{1,s}-d_{2,s}}{2}\right)\right)^\gamma}\right]$$
$$= E_t^{\mathbb{P}}[f(\alpha \cdot X_s) g(\beta \cdot X_s)], \qquad (37)$$

where

$$f(x) = e^x$$
$$g(x) = \frac{1}{(2\cosh(x))^\gamma}$$

and

$$\alpha = \left(1 - \frac{\gamma}{2}\right)\iota_1 - \frac{\gamma}{2}\iota_2, \quad \beta = \frac{\epsilon - 1}{2\epsilon}(\iota_1 - \iota_2).$$

Since X is affine and $g \in \mathcal{S}^*$, Theorem 1 readily applies to Equation (37). When the increments of X are i.i.d., the conditional characteristic function for X is known explicitly, which Martin (2011) uses to compute (37) following a Fourier transform for g.

Next, through the analytical results for the stock price in Equation (36), one can conveniently compute the volatility and conditional risk premium of the stock. To compute the risk premium, we can treat the stock as a portfolio of zero-coupon equities, each with a single dividend payment $D_{i,t}dt$. The risk premium of the stock is then the value-weighted average of the risk premium for these zero-coupon equities. For a zero-coupon equity, we can first compute its exposure to the primitive shocks in the state variable X_t and the risk prices associated with these shocks, which are given by their covariances with the stochastic discount factor m_t. Thus, by Itô's lemma, the conditional risk premium for any asset with price $P(X_t,t)$ is given by

$$E_t[R^e]=(\nabla_X \log m_t)'\sigma_t\sigma_t'(\nabla_X \log P_t), \tag{38}$$

where $\sigma_t\sigma_t'$ is the time-t covariance of the factors X_t. Here, $(\nabla_X \log m_t)'\sigma_t$ in (38) gives the price of risk for all the shocks in X_t. The term $(\nabla_X \log P_t)'\sigma_t$ gives the exposure that the asset has to these shocks. We may use our method to compute this sensitivity by integrating the sensitivities of the zero-coupon equity prices. In computing the zero-coupon equity sensitivity, the key step is computing the derivative of $H=E_t^{\mathbb{P}}[f(\alpha\cdot X_s)g(\beta\cdot X_t)]$ from Equation (37). Using the results of Theorem 1, we obtain that

$$\partial_j H=\frac{1}{2\pi}\langle\hat{g},\partial_j\psi(\alpha+\cdot\beta i)\rangle, \tag{39}$$

where from Equation (14) we have

$$\partial_j\psi(\alpha+is\beta)=B_j(T;\alpha+is\beta,\Theta)e^{A(T;\alpha+is\beta,\Theta)+B(T;\alpha+is\beta,\Theta)\cdot X_0}. \tag{40}$$

Here, B_j denotes the jth component of the vector B. Besides the equity premium, we can also compute the conditional volatility of stock return analytically using the above results on zero-coupon equity sensitivities to different shocks.

Several observations are in order. First, introducing additional state variables to X within the affine framework to capture richer dynamics of dividends (such as time-varying conditional moments of dividend growth) is quite straightforward. Due to the time-separable utility function, these additional state variables do not directly enter into the pricing equation (36) and thus will not result in a curse of dimensionality—the dimension of the problem remains exactly the same. Second, one can also further enrich the model by adding preference shocks that are *pl-linear* in the state variables, e.g., see the external habit models in Pástor and Veronesi (2005) or Bekaert and Engstrom (2010). See also the online appendix for a general specification. Third, we can extend the model to have more than two assets, which can be solved using the multidimensional version of the generalized transform in Proposition 2.

4.1 A calibrated example: Time-varying labor income risk

In this section, we use a general equilibrium model with time-varying labor income risk to illustrate how to apply the generalized transform method for

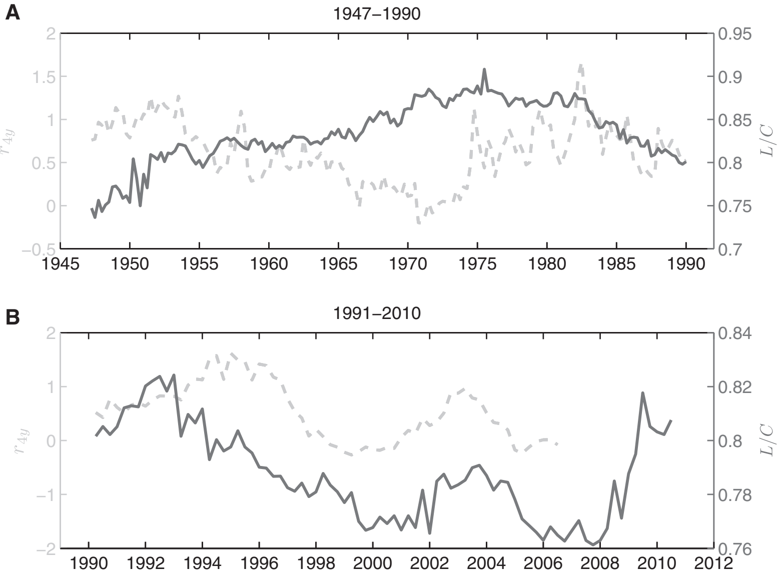

Figure 3
Long-term returns and labor share pre- and post-1990
Data are quarterly and the sample period is 1947Q1 to 2010Q2. L/C, share of labor income to consumption; r_{4y}, lagged four-year cumulative returns of the market portfolio.

economic modeling. The model also draws a number of new insights on how the time-varying covariance between labor income and dividends affects asset pricing. Santos and Veronesi (2006, hereafter SV) point out a natural source of time variation in risk premium via a *composition effect*: as the share of labor income in total consumption varies over time, so will the covariance between consumption and dividends, which in turn generates a time-varying equity premium. They illustrate this point in a model with stationary labor share and multiple financial assets, which provides very convenient closed-form solutions for asset prices.

We plot in Figure 3 the share of labor income and lagged four-year cumulative returns of the CRSP (Center for Research in Security Prices) value-weighted market index. We use per-capita consumption (nondurables and services) from the BEA and labor income series constructed following Lettau and Ludvigson (2001). Following SV, the labor share is defined as the ratio of labor income to consumption; thus, dividends are defined as the difference between consumption and labor income. Consistent with the findings of SV, the labor share and lagged market return in Panel A of Figure 3 are negatively correlated, with an average correlation of −0.35. However, in the post-1990 period, the two series become positively correlated, which is opposite of what the composition effect implies. These results suggest that other covariates may be playing a role in determining the relation between the labor share and

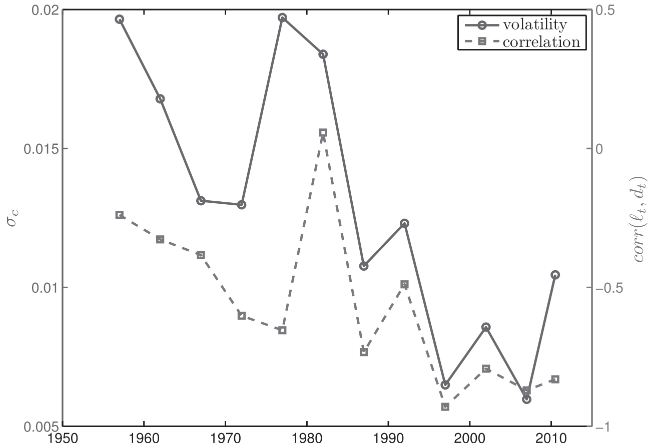

Figure 4
Consumption volatility and correlation between labor income and dividends
This figure plots the (annualized) standard deviation of aggregate consumption growth and the correlation between the growth rates of labor income and dividends in five-year windows. The data are quarterly from 1952Q1 to 2010Q2. At the end of each series, we add the estimate for the most recent five years.

the equity premium (see also Duffee 2005 and Kozhanov 2009 for related findings).

One example of such covariates is consumption volatility. Lettau, Ludvigson, and Wachter (2008) argue that declining macroeconomic volatility since the 1980s has played a key role in the decline of the equity premium. They estimate that a structural break occurred in consumption volatility in 1991, which motivates our choice of the two subsamples. In Figure 4, we plot the volatility of consumption growth for nonoverlapping five-year periods, using quarterly data from 1952 to 2010, and the correlation between labor income and dividends during the same periods. Interestingly, the correlation between labor income and dividends shows a pattern similar to volatility. All else equal, consumption volatility should indeed rise with the correlation between its two components. However, consumption volatility could also change independently of the correlation, e.g., through time-varying share of labor income in consumption or variations in the volatilities of labor income and dividends.

We propose a simple model that captures the dynamics of the conditional moments of labor income, dividends, and consumption suggested by the above evidence. The model is a special case of the two-asset models discussed previously, with dividends from the two assets interpreted as financial income (dividends) and labor income. We extend the models of Cochrane, Longstaff, and Santa-Clara (2008) and Martin (2011) in two important ways. We add a volatility factor, which simultaneously drives the conditional volatilities of labor income and dividends, as well as the correlation between the two.

Second, we impose cointegration between labor income and dividends so that the long-run labor share converges to a stationary distribution. Our model also differs from SV in that labor share and the correlation between labor income and dividends can move independently.

Specifically, let log dividends and log labor income be d_t and ℓ_t, and let V_t be a volatility factor. Suppose $X_t = (d_t, \ell_t, V_t)'$ follows an affine process

$$dX_t = \mu_t dt + \sigma_t dW_t. \tag{41}$$

We assume that the conditional drift μ_t is given by

$$\mu_t = \begin{pmatrix} \overline{g} - \quad a\kappa_s(\overline{s} - (\ell_t - d_t)) \\ \overline{g} + (1-a)\kappa_s(\overline{s} - (\ell_t - d_t)) \\ \kappa_V(\overline{V} - V_t) \end{pmatrix}. \tag{42}$$

This formulation gives the same average growth rate $\overline{g}$ for dividends and labor income. The second term in the growth rates of dividends and labor income allows for the shares of labor income and dividends to be stationary.[13] The parameter a gives us additional flexibility in specifying the degree of time variation in the growth rates of dividends and labor income. Equation (42) also implies that the volatility factor V_t is stationary with a long-run mean $\overline{V}$ and speed of mean reversion κ_V.

The conditional covariance of the factors is given by

$$\sigma_t \sigma_t' = \Sigma_0 + \Sigma_1 V_t, \tag{43}$$

where Σ_0 and Σ_1 can be any positive semidefinite matrices with the restriction $\Sigma_{0,33} = 0$ so that the volatility factor always remains positive. As V_t increases, the volatilities of dividends and labor income will increase. Moreover, the instantaneous correlation between dividends and labor income, ρ_t, also varies with the volatility factor V_t. The structure of a single volatility factor implies that the correlation and volatilities have to move in lockstep. Although this feature is clearly restrictive, it allows us to capture the comovement between consumption volatility and correlation demonstrated in Figure 4. One can substantially relax the covariance structure using the Wishart process, which allows for fully stochastic covariance between labor income and dividends. We calibrate the model to match moments for dividends and labor income in the data; for full details of model calibration, see Appendix 4. For simplicity, we assume that the shocks to the volatility factor are uncorrelated with dividend and labor income shocks.

We now examine the risk premium on financial wealth and human capital, which can be computed efficiently using Equations (38)–(40). The risk premium

[13] To see this, consider the dynamics of the log labor income-dividend ratio $s_t = \ell_t - d_t$. Equation (42) implies that the drift of s_t will be $\kappa_s(\overline{s} - s_t)$, with $\overline{s}$ being the long-run average of the log labor income-dividends ratio and $\kappa_s \geq 0$ the speed of mean reversion.

for financial wealth depends on its exposure to both dividend and labor income shocks. First, the financial wealth claim is directly exposed to dividend shocks via its cash flows (dividends). Second, via the discount factor, it is exposed to both dividend and labor income shocks. Both mechanisms will have an effect on the risk premium. For example, positive shocks to labor income will decrease the premium demanded for dividends, increase the risk-free rate (a higher share of the less volatile labor income tends to smooth consumption, reducing the precautionary savings motive), and increase expected future dividends due to the mean reversion in Equation (42) (with $a>0$). Our model incorporates all of these effects to determine the risk premium for the dividend claim. The value of the dividends tree will also be sensitive to covariance shocks (V_t), but these shocks are not priced in our model due to the assumption that they are not correlated with consumption.

Figure 5 plots the conditional risk premium on financial wealth and human capital as functions of the labor share and correlation. The plot focuses on the parameter region that is most relevant based on the stationary distributions of the two variables. When volatilities are high and the correlation is less negative, the model generates a significant composition effect. For example, when the conditional correlation between labor income and dividends $\rho_t = -0.1$, the conditional risk premium on financial wealth falls from 6.6% to 1.8% as labor share rises from 0.6 to 0.9. However, when $\rho_t = -0.8$, the risk premium essentially remains at zero for the same rise in labor share. Moreover, the risk premium on financial wealth is more sensitive to changes in volatility and correlation when labor share is low. As Figure 4 suggests, both consumption volatility and the correlation between labor income and dividends were indeed

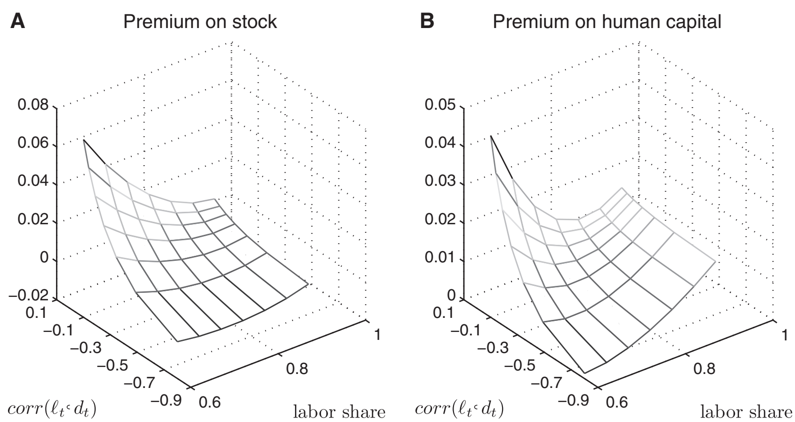

Figure 5
Conditional risk premium on financial wealth and human capital
The left panel plots the conditional risk premium on the stock (financial wealth), and the right panel plots the conditional risk premium on human capital as a function of labor share L_t/C_t and correlation ρ_t.

higher pre-1990 and drop significantly in the following period. Thus, our model could potentially explain the "disappearing" composition effect in the data.[14]

Through the lens of our model, we can also analyze the comovement between the risk premium on financial wealth and human wealth. Previous studies have different findings when measuring the sign of this comovement in the data. For example, see Hansen et al. (2007) and Lustig and Van Nieuwerburgh (2008). Cash flows from the claim on financial and human wealth are negatively correlated most of the time in our model. However, both positive and negative correlation between the risk premium on financial and human wealth can occur. The risk premium on financial wealth and human wealth will be negatively correlated when the composition effect is the main driver of variations in risk premium over time. However, when changes in the volatilities and correlation become the main driver of variations in risk premium, the two risk premiums can become positively correlated.

This example illustrates the power of our method to generate new economic insights through the analysis of general equilibrium models. The general class of models that we study here of heterogeneous agents and heterogeneous goods can also be used to examine the general equilibrium effects of the cross-section of stocks as well as international finance models. Our method makes many of these models tractable for analysis.

5. Conclusion

We provide analytical results for computing 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. Our method can also be applied to other processes with tractable characteristic functions, such as discrete-time affine processes, Lévy processes, and Markov-switching affine processes.

We demonstrate the power of this method with two examples. First, we study the pricing of defaultable bonds with stochastic recovery. We show that not only does the commonly used constant recovery assumption lead to substantial pricing errors in comparison to the stochastic recovery model, but the latter exhibits important nonlinearity that cannot be replicated by constant recovery models. In the second example, we apply the generalized transform method in a general equilibrium model of time-varying labor income risk. The model not only helps explain the changing predictive power of labor share with declining volatility but also shows that the risk premium on financial wealth

[14] For further details of the mechanism, see the online appendix.

and human capital can be positively or negatively correlated, depending on whether variations in labor share or covariances are the main driver of risk premium.

Appendix

1. Proof of Theorem 1

$$\dot{B}=K_1^\top B+\frac{1}{2}B^\top H_1 B-\rho_1+\ell_1(\phi(B)-1) \quad B(0)=\alpha+is\beta, \tag{A1}$$

$$\dot{A}=K_0^\top B+\frac{1}{2}B^\top H_0 B-\rho_0+\ell_0(\phi(B)-1) \quad A(0)=0, \tag{A2}$$

where $\phi(c)=E_\nu[e^{c\cdot Z}]$, the moment-generating function of the jump distribution and $(B^\top H_1 B)_k=\sum_{i,j} B_i H_{1,ijk} B_k$. Solving the ODE system (A1–A2) adds little complication to the transform. The solution is available in closed form in some cases and generally can be quickly and accurately computed using standard numerical methods.

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 (A1)–(A2).

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

$$\frac{dF}{dP}=\frac{e^{-\int_0^T r_\tau d\tau}e^{\alpha\cdot X_T}}{E_0[e^{-\int_0^T r_\tau d\tau}e^{\alpha\cdot X_T}]}, \tag{A3}$$

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,

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

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 typically can 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,

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

In the last equation, we interpret $g \in \mathcal{S}^*$. By Assumption 2, $f^F_{\beta \cdot X_T} \in \mathcal{S}$ and so $\hat{f}^F_{\beta \cdot X_T} \in \mathcal{S}$ also. Thus, Fourier inversion holds and $(\hat{f}^F_{\beta \cdot X_T})\check{} = f^F_{\beta \cdot X_T}$ (see Corollary 8.28 in Folland 1984), where we denote the inverse Fourier transform of a function h by $\check{h}(s) = \frac{1}{2\pi}\int e^{is\cdot x}h(x)dx$. Applying this,

$$H = F_0 \; \langle g, (\hat{f}^F_{\beta \cdot X_T})\check{}\rangle$$

$$= F_0 \; \langle \hat{g}, \check{f}^F_{\beta \cdot X_T}\rangle.$$

This equation holds because of Fourier inversion and the definition of Fourier transform of a tempered distribution. Notice that when both $f_{\beta \cdot X_T}$ and g are in $\mathcal{S}$, we can write this last equality as

$$\frac{F_0}{2\pi}\int_x g(x)\int_s \hat{f}^F_{\beta \cdot X_T} e^{isx} ds dx = \frac{F_0}{2\pi}\int_s \hat{f}^F_{\beta \cdot X_T}\int_x g(x)e^{isx} dx ds.$$

Thus, we see that the theory from Fourier analysis of tempered distributions justifies the change of order of integration in a general sense. We can therefore further simply obtain

$$H = \langle \hat{g}, F_0 \check{f}^F_{\beta \cdot X_T}\rangle$$

$$= \frac{1}{2\pi}\langle \hat{g}, \psi(\alpha + \cdot \beta i)\rangle.$$

This last step holds by Assumption 1. This is the desired result. □

In some cases of interest, Assumption 2 may be violated. It could be that $\beta \cdot X_T$ has heavy tails so that, e.g., $E[(\beta \cdot X_T)^4] = \infty$. 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 nonsmooth densities).

2. Proof of Proposition 1

In analogy to Duffie, Pan, and Singleton (2000) and Pan (2002), define

$$G(\alpha_0; v, n|x, t) = e^{A_t + B_t \cdot x}\sum_{|\xi|=n}\binom{n}{\xi}L(x)^{\xi}, \tag{A4}$$

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

$$\dot{B} = K_1^\top B + \frac{1}{2}B^\top H_1 B - \rho_1 + \lambda_1(\phi(B) - 1), \quad B(0) = \alpha_0, \tag{A5}$$

$$\dot{A} = K_0^\top B + \frac{1}{2}B^\top H_0 B - \rho_0 + \lambda_0(\phi(B) - 1), \quad A(0) = 0, \tag{A6}$$

$$\partial_1 \dot{B} = K_1^\top \partial_1 B + \partial_1 B^\top H_1 B + \lambda_1 \nabla\phi(B) \cdot \partial_1 B, \quad \partial_1 B(0) = v, \tag{A7}$$

$$\partial_1 \dot{A} = K_0^\top \partial_1 B + \partial_1 B^\top H_0 B + \lambda_0 \nabla\phi(B) \cdot \partial_1 B, \quad \partial_1 A(0) = 0, \tag{A8}$$

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

$$\partial_m \dot{B} = K_1^\top \partial_1 B + \frac{1}{2}\sum_{i=0}^{m}\binom{m}{i}\partial_i B^\top H_1 \partial_{m-i} B + \partial_{m-1}(\lambda_1 \nabla\phi(B) \cdot \partial_1 B), \; \partial_m B(0) = 0, \tag{A9}$$

$$\partial_m \dot{A} = K_0^\top \partial_1 B + \frac{1}{2}\sum_{i=0}^{m}\binom{m}{i}\partial_i B^\top H_1 \partial_{m-i} B + \partial_{m-1}(\lambda_0 \nabla\phi(B) \cdot \partial_1 B), \; \partial_m A(0) = 0. \tag{A10}$$

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 (A1) for any initial condition of the form $\alpha_0 = \alpha + is\beta$ for any $s \in \mathbb{R}$. Additionally, for any such 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 $\Psi_t^n(i, x) = e^{-i} G(\alpha, v, n | x, T-t)$ 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,
   $$\frac{dF}{dP} = \frac{e^{-\int_0^T r_\tau d\tau} e^{\alpha \cdot X_T} (v \cdot X_T)^n}{E_0[e^{-\int_0^T r_\tau d\tau} e^{\alpha \cdot X_T} (v \cdot X_t)^n]}, \tag{A11}$$
   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.

3. Generalized Transform in Discrete Time

In this appendix, we show how our method applies in a discrete-time setting. Here, we replace (7) with

$$\Delta X_t = (K_0 + K_1 X_t) + \epsilon_{t+1}, \tag{A12}$$

where ϵ_{t+1} has a conditional distribution, which depends on X_t, which satisfies

$$E\left[e^{\alpha \cdot \epsilon_{t+1}}\right] = e^{\hat{A}(\alpha) + \hat{B}(\alpha) \cdot X_t}. \tag{A13}$$

For example, if $\hat{A}(\alpha) = \frac{1}{2}\alpha' \Sigma \alpha$ and $\hat{B}(\alpha) = 0$, it follows that $\epsilon_t \sim N(0, \Sigma)$, in which case X_t follows a simple vector autoregression. In general, the discrete-time family of processes given by (A12–A13) is quite flexible and allows for jump-type processes and time-varying covariance (see, e.g., Le, Singleton, and Dai 2010).

For the discrete-time processes, the analogous version of Theorem 1 gives

$$H_D(g, \alpha, \beta) = E_0\left[\exp\left(-\sum_{u=0}^{T} r(X_u)\right) e^{\alpha \cdot X_T} g(\beta \cdot X_T)\right]$$

$$= \frac{1}{2\pi} \langle \hat{g}, \psi(\alpha + \cdot \beta i) \rangle, \tag{A14}$$

where now (A1–A2) are replaced by

$$\Delta B = K_1^\top B + \hat{B}(B) - \rho_1, \quad B(0) = \alpha + is\beta, \tag{A15}$$

$$\Delta A = K_0^\top B + \hat{A}(B) - \rho_0, \quad A(0) = 0. \tag{A16}$$

4. Labor Income Risk

This section provides more details on the calibration and analysis of the model with time-varying labor income risk.

The parameters are summarized in Table A1 and are calibrated as follows. We set the long-run mean growth rate of labor income and dividends to 1.5%. We specify the long-run labor income share, $\bar{S}$, to be 75%. As the covariance parameters (Σ_0 and Σ_1) are difficult to directly interpret, we calibrate them by considering their effect on the volatility of labor income, the

Table A1
Parameters

Parameter	Value	Parameter	Value
γ	6	$\bar{\rho}_{\ell,d}$	-30.3%
ρ	1%	$\bar{\sigma}_{\ell}$	5.4%
$\overline{g}$	1.5%	$\bar{\sigma}_{d}$	11.1%
$\overline{S}$	80%	$\sigma_{\infty}(\rho_{\ell,d})$	9.8%
a	$\frac{1}{3}$	$\sigma_{\infty}(\sigma_{\ell})$	0.0018%
κ_s	0.0231	$\sigma_{\infty}(\sigma_{d})$	0.017%
κ_V	0.0693	$\sigma_{SS}(V)$	1.07

This table gives the parameters and moments used to calibrate the model. The left column gives the preference parameters and conditional mean parameters for the process. The right column gives the conditional moments used to calibrate the parameters (Σ_0, Σ_1). The first three calibration moments refer to the steady-state values. The next three refer to the conditional volatility of the conditional moments evaluated at the long-run mean of V. $\sigma(\sigma_d)$ is the steady-state volatility of σ_d. $\bar{V}$ is normalized to one.

volatility of dividends, and their correlation. We set the parameters so that when V_t is at its long-run mean $\bar{V}$ (which is normalized to be one), $(\sigma_{\ell,t}, \sigma_{d,t}, \rho_{\ell,d,t})$ are given by $\bar{\sigma}_{\ell} = 5.4\%$, $\bar{\sigma}_d = 11.1\%$ and $\bar{\rho}_{\ell,d} = -30.3\%$, respectively. Note that due to CRRA utility, our model presents the equity premium puzzle (Mehra and Prescott 1983), and we choose our parameterization to generate a higher premium with reasonable risk aversion by slightly overstating the volatility of labor income relative to the data, with the ratio of dividend to labor income volatility qualitatively similar to Lettau, Ludvigson, and Wachter (2008). We also calibrate the volatility of $(\sigma_d, \sigma_{\ell}, \rho_{\ell,d})$ when V_t is at its long-run mean, which we denote with by $\sigma_{\infty}(\sigma_d) = 1.7$bp, $\sigma_{\infty}(\sigma_{\ell}) = 0.18$bp, and $\sigma_{\infty}(\rho_{\ell,d}) = 9.8\%$. Finally, we calibrate the volatility of V in the steady-state distribution, which we denote by $\sigma_{SS}(V)$, to be 1.07. Taken together, these seven moments (along with the simplifying assumption that innovations to V are uncorrelated with innovations to either ℓ or d) fix the free parameters in Σ_0 (three parameters) and Σ_1 (four parameters). Under this calibration, when V is at the highest (lowest) decile, the volatility of labor income is 6% (5%), the volatility of dividends is 16% (6%), and their correlation is -10% (-80%). The volatility parameters were chosen to qualitatively match the variation found in Figure 4.

References

Abramowitz, M., and I. A. Stegun. 1964. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover Books.

Acharya, V. V., S. T. Bharath, and A. Srinivasan. 2007. Does Industry-wide Distress Affect Defaulted Firms? Evidence from Creditor Recoveries. *Journal of Financial Economics* 85:787–821.

Altman, E. I., B. Brady, A. Resti, and A. Sironi. 2005. The Link Between Default and Recovery Rates: Theory, Empirical Evidence, and Implications. *Journal of Business* 78:2203–28.

Altman, E. I., and B. J. Kuehne. 2011. Defaults and Returns in the High-yield Bond and Distressed Debt Market: The Year 2010 in Review and Outlook. New York: New York University Salomon Center.

Ang, A., G. Bekaert, and M. Wei. 2008. The Term Structure of Real Rates and Expected Inflation. *Journal of Finance* 63:797–849.

Bakshi, G., C. Cao, and Z. Chen. 1997. Empirical Performance of Alternative Option Pricing Models. *Journal of Finance* 52:2003–49.

Bakshi, G. S., and Z. Chen. 1997. An Alternative Valuation Model for Contingent Claims. *Journal of Financial Economics* 44:123–65.

Bakshi, G., and D. Madan. 2000. Spanning and Derivative-security Valuation. *Journal of Financial Economics* 55:205–38.

Bakshi, G., D. Madan, and F. Zhang. 2006. Understanding the Role of Recovery in Default Risk Models: Empirical Comparisons and Implied Recovery Rates. Working Paper, University of Maryland.

Bates, D. S. 1996. Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options. *Review of Financial Studies* 9:69–107.

Bekaert, G., and E. Engstrom. 2010. Asset Return Dynamics Under Bad Environment-Good Environment Fundamentals. Working Paper, Columbia University.

Buraschi, A., F. Trojani, and A. Vedolin. 2010. Economic Uncertainty, Disagreement, and Credit Markets. Working Paper, Imperial College Business School.

Carr, P., and D. B. Madan. 1999. Option Valuation Using the Fast Fourier Transform. *Journal of Computational Finance* 2:61–73.

Carr, P., H. Geman, D. B. Madan, and M. Yor. 2002. The Fine Structure of Asset Returns: An Empirical Investigation. *Journal of Business* 75:305–32.

Chacko, G., and S. Das. 2002. Pricing Interest Rate Derivatives: A General Approach. *Review of Financial Studies* 15:195–241.

Chen, R.-R., and L. Scott. 1995. Interest Rate Options in Multifactor Cox-Ingersoll-Ross Models of the Term Structure. *Journal of Derivatives* 3:53–72.

Chen, H. 2010. Macroeconomic Conditions and the Puzzles of Credit Spreads and Capital Structure. *Journal of Finance* 65:2171–212.

Chen, H., and S. Joslin. 2011. Online Appendix to Generalized Transform Analysis of Affine Processes and Applications in Finance. Working Paper, University of Southern California.

Chen, H., S. Joslin, and N.-K. Tran. 2010. Affine Disagreement and Asset Pricing. *American Economic Review* 100:522–26.

Chen, H., S. Joslin, and N.-K. Tran. Forthcoming. Rare Disasters and Risk Sharing with Heterogeneous Beliefs. *Review of Financial Studies*.

Cochrane, J. H., F. Longstaff, and P. Santa-Clara. 2008. Two Trees. *Review of Financial Studies*. 21:347–85.

Collin-Dufresne, P., R. Goldstein, and J. Hugonnier. 2004. A General Formula for Valuing Defaultable Securities. *Econometrica* 72:1377–407.

Dai, Q., K. J. Singleton, and W. Yang. 2007. Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bond Yields. *Review of Financial Studies* 20:1669–706.

Duffee, G. R. 1999. Estimating the Price of Default Risk. *Review of Financial Studies* 12:197–226.

Duffee, G. R. 2005. Time Variation in the Covariance Between Stock Returns and Consumption Growth. *Journal of Finance* 60:1673–712.

Duffie, J. D. 2005. Credit Risk Modeling with Affine Processes. *Journal of Banking and Finance* 29: 2751–802.

Duffie, J. D., D. Filipovic, and W. Schachermayer. 2003. Affine Processes and Applications in Finance. *Annals of Applied Probability* 13:984–1053.

Duffie, J. D., J. Pan, and K. Singleton. 2000. Transform Analysis and Asset Pricing for Affine Jump-diffusions. *Econometrica* 68:1343–76.

Duffie, J. D., M. Schroder, and C. Skiadas. 1996. Recursive Valuation of Defaultable Securities and the Timing of Resolution of Uncertainty. *Annals of Applied Probability*. 6:1075–90.

Duffie, J. D., and K. J. Singleton. 1999. Modeling Term Structures of Defaultable Bonds. *Review of Financial Studies* 12:687–720.

Dumas, B., A. Kurshev, and R. Uppal. 2009. Equilibrium Portfolio Strategies in the Presence of Sentiment Risk and Excess Volatility. *Journal of Finance* 64:579–629.

Folland, G. 1984. Real Analysis: Modern Techniques and Their Applications. Hoboken, NJ: Wiley-Interscience.

Gabaix, X. 2009. Linearity-generating Processes: A Modeling Tool Yielding Closed Forms for Asset Prices. Working Paper, New York University.

Hansen, L. P., J. Heaton, J. Lee, and N. Roussanov. 2007. Intertemporal Substitution and Risk Aversion. In J. J. Heckman, and E. E. Leamer (eds.), *Handbook of Econometrics*, vol. 6, pp. 3967–4056. Amsterdam: Elsevier.

Heston, S. L. 1993. A Closed-form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*. 6:327–43.

Joslin, S. 2010. Pricing and Hedging Volatility in Fixed Income Markets. Working Paper, University of Southern California.

Kozhanov, I. 2009. The Challenge of the Composition Effect. Working Paper, State University of New York at Buffalo.

Le, A., K. Singleton, and Q. Dai. 2010. Discrete-time Affine$^{\mathbf{Q}}$ Term Structure Models with Generalized Market Prices of Risk. *Review of Financial Studies* 23:2184–227.

Lettau, M., and S. Ludvigson. 2001. Consumption, Aggregate Wealth, and Expected Stock Returns. *Journal of Finance*. 56:815–49.

Lettau, M., S. C. Ludvigson, and J. A. Wachter. 2008. The Declining Equity Premium: What Role Does Macroeconomic Risk Play? *Review of Financial Studies* 21:1653–87.

Lustig, H., and S. Van Nieuwerburgh. 2008. The Returns on Human Capital: Good News on Wall Street Is Bad News on Main Street. *Review of Financial Studies* 21:2097–137.

Madan, D. B., P. P. Carr, and E. C. Chang. 1998. The Variance Gamma Process and Option Pricing. *European Finance Review* 2:79–105.

Martin, I. 2011. The Lucas Orchard. Working Paper, Stanford University.

Mehra, R., and E. C. Prescott. 1983. The Equity Premium: A Puzzle. *Journal of Monetary Economics* 15: 145–61.

Menzly, L., T. Santos, and P. Veronesi. 2004. Understanding Predictability. *Journal of Political Economy* 112:1–47.

Pan, J. 2002. The Jump-risk Premia Implicit in Options: Evidence from an Integrated Time-series Study. *Journal of Financial Economics* 63:3–50.

Pan, J., and K. J. Singleton. 2008. Default and Recovery Implicit in the Term Structure of Sovereign “CDS” Spreads. *Journal of Finance* 63:2345–384.

Pástor, L., and P. Veronesi. 2005. Rational IPO Waves. *Journal of Finance* 60:1713–757.

Piazzesi, M. 2005. Bond Yields and the Federal Reserve. *Journal of Political Economy* 113:311–44.

Piazzesi, M., M. Schneider, and S. Tuzel. 2007. Housing, Consumption, and Asset Pricing. *Journal of Financial Economics* 83:531–69.

Protter, P. E. 2004. Stochastic Integration and Differential Equations. New York: Springer.

Santos, T., and P. Veronesi. 2006. Labor Income and Predictable Stock Returns. *Review of Financial Studies*. 19:1–44.

Singleton, K. J. 2001. Estimation of Affine Asset Pricing Models Using the Empirical Characteristic Function. *Journal of Econometrics* 102:111–41.