Generalized Transform Analysis of Affine Processes And Asset Pricing Applications[*]

Hui Chen[†]   Scott Joslin[‡]

February 17, 2009

Abstract

Non-linearities are often important considerations in many applications in finance and economics such as pricing securities, computing equilibria, and conducting structural estimations. This paper generalizes the class of transform analysis for affine jump diffusions in Duffie, Pan, and Singleton (2000), which facilitates analytical treatment of a wide range of such problems. We apply our method to several examples, including option pricing, modeling nonlinear Taylor rules, modeling correlated defaults, and GMM estimation. We also provide detailed analysis of the power of our method in three concrete examples: (1) pricing defaultable bonds with state-dependent recovery; (2) computing the equilibrium of a Lucas economy with non-i.i.d. trees; and (3) computing equilibrium in an economy with investors having heterogeneous beliefs about a variety of fundamentals .

1

1 Introduction

In this paper, we provide analytical treatment of a class of transforms for state variables that follow affine jump-diffusions (AJD). With the ability to handle a very general class of nonlinearities, the transform analysis brings analytical and computational tractability to a wide range of problems in economics and finance. We illustrate the application of the generalized transform analysis with a variety of examples, including pricing options, term structure modeling, credit risk modeling, method of moment estimations, computing the equilibrium of consumption-based asset pricing models, and a general class of difference-of-belief models.

For a state variable X_t that follows an affine process (in the sense that the conditional characteristic function is affine[1]), Duffie et al. (2000), hereafter DPS, derive closed-form expression for the following transform:

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

where $R\left(X\right)$ is an affine function of X, which can be interpreted as the stochastic "discount rate", and $e^{u \cdot X_T} \left(v_0 + v_1 \cdot X_T\right) 1_{\{\alpha \cdot X_T < y\}}$ 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 generalized transform:

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

where f can be linear function, 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 $f(X) = e^{u \cdot X} \left(v_0 + v_1 \cdot X\right)$ and $g(\alpha \cdot X) = 1_{\{\alpha \cdot X_T < y\}}$, we recover the transform of DPS. The abundant flexibility in choosing function g helps address many nonlinearity problems in pricing (nonlinear discount factors or payoffs), estimation (nonlinear moments), and economic modeling. We provide several example applications for the generalized transform analysis to highlight its power.

The primary analytic tool that we use is the Fourier transform. In particular, we utilize knowledge of the conditional characteristic function of the state variabel X_t (under certain forward measures) jointly with a Fourier decomposition of the non-linearity in g. As we show, this combination produces expressions for our generalized transform which remain tractable by avoiding intermediate Fourier inversions.

Option pricing. When pricing standard European options, the payoff function can be expressed as the product of an exponential function and an indicator function. In this case, we recover the DPS transform as a special case.

---

[1] See Duffie et al. (2003) for an elaboration on the characterization via the characteristic function.

Term structure modeling. We consider a nonlinear Taylor Rule model that generalizes the model of fed funds target in Piazzesi (2005). With the generalized transform, we have more freedom in modeling the fed policy function. The policy function can be chosen to maintain the requirement that the fed target rate f is non-negative, has increments in a multiple of 25 basis points, and depend on macro variables such as GDP growth and inflation, but without the restriction that the distribution of f have the exponential-affine Laplace transform. One application of such a model is to price fed funds futures.

Credit risk modeling. We introduce a simple model of default contagion that violates the doubly stochastic assumption in that the default event of one firm affects the default probability of another firm beyond its own default intensity. The flexibility of modeling the default probability for the secondary firm after the default of the primary firm makes it easy to capture the nonlinearity and time-dependence of the contagion effects.

Defaultable bond is a good example with nonlinear payoffs, since the recovery rate at default has to be between 0 and 1, and has been empirically shown to nonlinearly depend on macro and firm-specific variables. We introduce a class of state-dependent recovery models into the reduced-form models of default, which substantially relaxes the "recovery of market value" assumption standard in current literature. We derive closed form solutions for the pricing of defaultable zero-coupon bond. The model can also be used to price other credit derivatives such as credit default swaps or recovery locks[2]. Our example of Cauchy recovery model demonstrates that ignoring the correlation between recovery rates and default intensity can lead to substantial deviations in credit spreads, especially for bonds with high or low credit quality.

Method of moments estimation. The need to compute unconditional and conditional moments of nonlinear functions arise in the method of moments and GMM estimations. We use the example of term structures of conditional corporate default probabilities and provide a GMM estimation procedure.

Computing the equilibrium of asset pricing models. Besides nonlinear payoffs, nonlinear discount factors often arise in asset pricing models. We use the example of the Lucas model with two trees. We show that it is straightforward to extend the two tree model of Cochrane et al. (2008) to allow for mean reversion and conditional heteroscedasticity in dividend growth. We use the model to analyze the equilibrium effect of growth-value stocks.

Difference-of-opinion models. Even when fundamentals follows simple dynamics, non-linearity in discount factors may arrive due to heterogeneous beliefs among agents. Agents may disagree regarding expected changes and volatility of fundamentals and also the likelihood and magnitude of jumps in the state of the economy. In equilibrium, beliefs then become important determinants

---

[2]This 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.

of asset prices. Thus, differences in opinions regarding higher order beliefs become relevant as well. We examine the effects of disagreement regarding jump intensity on asset prices.

Our paper is most closely related to Duffie et al. (2000), which is widely used in term structure modeling, option pricing, and econometric estimations of affine models. Examples include Singleton (2001), Pan (2002), Piazzesi (2005), and Joslin (2009), etc. The generalized transform in this paper extends the DPS method to a general class of nonlinear functions.

Fourier transform methods are commonly used in the solutions of nonlinear PDEs which arise natural in finance and economics through the Feynman-Kac methodology. For example, Heston (1993) computes the risk-neutral exercise probabilities for options in an affine stochastic volatility model through Fourier inversion of the conditional characteristic function. Martin (2008) uses the Fourier transform to address the nonlinearities in the pricing kernel that arise in a Lucas economy with two *i.i.d.* endowment processes. The conditional characteristic functions from the Fourier transform is known in closed form in this special case. Using the generalized transform, we can not only relax the *i.i.d.* restriction and model log dividends as a general affine process instead , but also allow for different utility functions and preference shocks.

The paper proceeds as follows. Section 2 gives the main theoretical results on the transform. Section 3 outlines various areas where the transform analysis can be applied. Section 4 investigates three concrete example applications of the generalized transform analysis. Section 5 concludes.

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 $(\Omega, \mathcal{F}, P)$ and an information filtration $\{\mathcal{F}_t\}$, and suppose that X is a Markov process in some state space $D \subset \mathbb{R}^n$ satisfying the stochastic differential equation

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

where B is an $\mathcal{F}_t$-standard n-dimensional Brownian motion and Z is pure jump process with arrival intensity $\lambda_t = \lambda_0 + \lambda_1 \cdot X_t$ with fixed D-invariant distribution ν. For brevity, let Θ denote the parameters of the process $(K_0, K_1, H_0, H_1, \lambda_0, \lambda_1, \nu)$. Alternatively, we can more precisely define the process in terms of the infinitesimal generator or, as Duffie et al. (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 $f : \mathbb{R}^N \to \mathbb{R}$ 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 piecewise 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

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

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:

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

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*

$$H(f, g, \alpha, \beta) = E_0 \left[\exp\left(-\int_0^T R(X_s) ds \right) f(\alpha \cdot X_T) g(\beta \cdot X_T) \right]$$
$$= \langle \hat{g}, G(\alpha + \cdot \beta i) \rangle \tag{6}$$

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

$$s \mapsto G(\alpha + s\beta i) = E_0[e^{-\int_0^T r_u du} e^{\alpha + is\beta}]. \tag{7}$$

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 ODE/BVP:

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

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

where $\phi(c) = E_\nu[e^{c \cdot Z}]$, the moment-generating function of the jump distribution.

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

$$H = \int_{-\infty}^{\infty} \hat{g}(s) G(\alpha + is\beta) ds \tag{10}$$

5

Fourier transforms of many functions are known in closed form. See, for example, Folland (1984). Additionally, standard rules allow for differentiation, integrations, products, convolutions and other operations to be conducted while maintaining closed form expressions. However, even in the case that the function $\hat{g}$ is not known in closed form, it can be computed readily.

In some cases of interests, Assumption 2 may be violated. $\beta \cdot X_T$ may have 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.)

It is worth noting that there is some flexibility in the choice of α and g. This is because $e^{\alpha \cdot X_T} g(\beta \cdot X_T) = e^{(\alpha - c\beta) \cdot X_T} \tilde{g}(\beta \cdot X_T)$ where $\tilde{g}(s) = e^{cs} g(s)$. This flexibility may be beneficial in the case that g in not integrable but decreases rapidly as s approaches either $+\infty$ or $-\infty$.[3] In this case, it may be computationally beneficial use such a transformation to apply (10) and avoid the use of distributions.

2.2 Extensions of Generalized Transforms

The result above is easily extended in a number of ways. First, we can extend f to exponential polynomials. That is, f can take the form

$$f(\alpha, v, X) = \sum_i p_i(v_i \cdot X) \operatorname{Real}\left(e^{\alpha_i \cdot X}\right), \tag{11}$$

where $\{p_i\}$ are arbitrary polynomials and $\{\alpha_i\}$ are complex vectors.[4] To stress the flexibility of this form, we will refer to any function which can be expressed as the product of a polynomial and a log-linear function as in (11) as *pl-linear* (polynomial-log-linear), while the rest of the functions as *non-pl-linear*.

Proposition 1. *Suppose that $v_i, \alpha_i \in \mathbb{R}^N$, $g \in \mathcal{S}^*$ and (Θ, α, β) satisfies Assumption 1 and Assumption 2 in Appendix A. Then*

$$H(f, g, \alpha, \beta) = E_0\left[\exp\left(-\int_0^T R(X_s)ds\right)(v_i \cdot X_T)^{n_i} e^{\alpha_i \cdot X_T} g(\beta \cdot X_T)\right]$$
$$= \langle \hat{g}, G(\alpha + \cdot \beta i)\rangle \tag{12}$$

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

$$s \mapsto G_i(\alpha + s\beta i) = E_0[e^{-\int_0^T r_u du}(v_i \cdot X_T)^{n_i} e^{\alpha + is\beta}]. \tag{13}$$

The function G_i is computed by solving the associated ODE/BVP in Appendix A.

---

[3]The logit function is such an example.

[4]Allowing complex eigenvalues allows one to have oscillatory sin and cos terms.

Next, at the expense of a multi-dimensional Fourier inversion, we can allow g to depend on X in more general ways. That is, instead of $g(\beta \cdot X)$, we consider $g(\beta_1 \cdot X, \cdots, \beta_M \cdot X)$ for $M \in \mathbb{N}$.

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

$$H(f, g, \alpha, \beta) = E_0 \left[\exp\left(-\int_0^T R(X_s)ds \right) f(\alpha \cdot X_T) g(BX_T) \right]$$
$$= \langle \hat{g}, G(\alpha + \cdot \beta i) \rangle \tag{14}$$

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

$$G_M : \mathbb{C}^M \to \mathbb{C} \ , \ s \mapsto G_M(\alpha + s^\top B i) = E_0[e^{-\int_0^T r_u du} e^{\alpha + i s^\top B}]. \tag{15}$$

This extension relaxes an important restriction for the nonlinear function $g(\cdot)$ in the generalized transform in Theorem 1, where $g(\cdot)$ only depends on X through the linear combination $\beta \cdot X$. In that case, the marginal impact of X_i on g will be proportional to β_i, which might be too restrictive in some cases. Proposition 2 removes this restriction, with the caveat that the problem is subject to the curse of dimensionality as M gets large.

3 Applications

3.1 Option Pricing

We first show that the generalized transform analysis can recover the DPS result as a special case. We use option pricing as an example.

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

$$M = E_t^Q \left[e^{-\int_0^T r_s ds + \alpha \cdot X_T} g_y(\beta \cdot X_T) \right]. \tag{16}$$

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, $y = \log K$,

$$P_0 = E_t^Q \left[e^{-\int_0^T r_s ds + y} g_y(X_T) \right] - E_t^Q \left[e^{-\int_0^T r_s ds + X_T} g_y(X_T) \right]. \tag{17}$$

However, the fourier transform is defined as a distribution:

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

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

$$\int \delta(x)h(x)dx = h(0) \tag{19}$$

7

This replicates the formula given in DPS obtained by Levy-inversion.

3.2 Nonlinear Taylor Rule

Through open market operations, the federal reserve targets the federal funds rate, the interest rate at which depository institutions lend balances at the Federal Reserve to other depository institutions overnight. The target is set by the Federal Open Market Committee which holds eight regularly scheduled meeting throughout the year as well as additional meetings whenever needed. The federal funds target rate can be viewed as a primitive in determining the yield curve.

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

$$f_t = \beta_0 + \beta_1 \pi_t + \beta_2 g_t + \epsilon_t \tag{20}$$

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.

Several potential deficiencies with such a simple rule are apparent. The fed target rate necessarily must be non-negative. This directly induces a nonlinearity. Additionally, the fed must implicitly consider this lower bound for future policy in setting current policy. Also, the target is typical a multiple of 25 basis points. Finally, the policy rule may incorporate other variables such as credit conditions or forward-looking variables.

Piazzesi (2005) uses pure jump processes with deterministic jump times to model 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, as well as possibly other macroeconomic variables such as unemployment. Piazzesi (2005) computes the prices of fixed income securities in this setting when the timing of jumps in f occur with stochastic intensity during FOMC meetings. In this specification, the distribution of 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

$$E_{t_0}[e^{af_t}] = e^{A_a + B_a \cdot X_{t_0}}. \tag{21}$$

This assumption implies, among other things, that the expected policy response is linear in the state. In other words, in expecation 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. However, as elaborated below, our specification allows for a nonlinear policy response and thus inherently nonlinear expectations. We elaborate below on the empirical relevance of such nonlinearities.

Figure 1 plots the historical quarterly time series of the fed funds rate. Also plotted are the fed
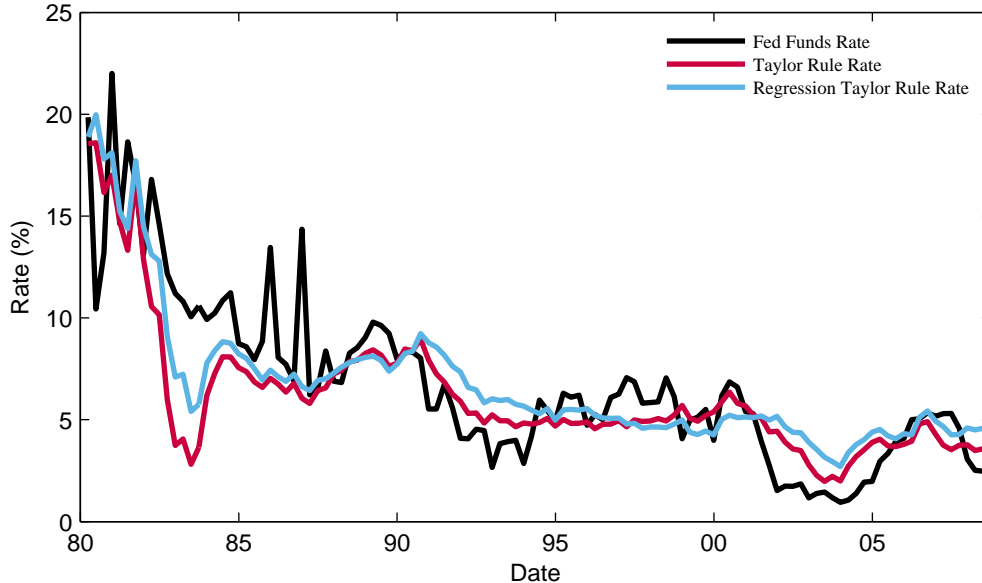
Figure 1: **Federal Funds Rate and Taylor Rule.** This figure plots the Federal Funds Rate on a quarterly basis. The red line indicates the rate prescribed by the simple Taylor rule (22). The blue line indicates the rate prescribed by the regression Taylor rule (23).

funds rate prescribed by two Taylor Rules:

$$r_t = 1\% + 1.5\pi_t + .5g_t + \epsilon_t \tag{22}$$

$$r_t = 1.19\% + 1.49\pi_t + .09g_t + \epsilon_t \tag{23}$$

The parameters in (22) correspond to the parameters in Taylor (1993) while the parameters in (23) are computed by an OLS regression. The federal funds rate is from Federal Reserve release H.15. The output gap is the real GDP output gap as computed by the Congressional Budget Office.

For both cases, the rule underestimates the response when actual rates are lowest (see Figure 2.) This suggests a nonlinearity in the form of more aggressive action in extreme situations. However, the natural lower limit of 0% also limits the policy action in extreme situations, introducing another nonlinearity. Thus it may be very natural to generate a simple *nonlinear* policy rule to model the Fed's policy action.[5]

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

---

[5]Another approach, of course, would be to have a full economic model of optimal policy rather than a simple rule. As we see empirically, such a rule is likely to be nonlinear as well.
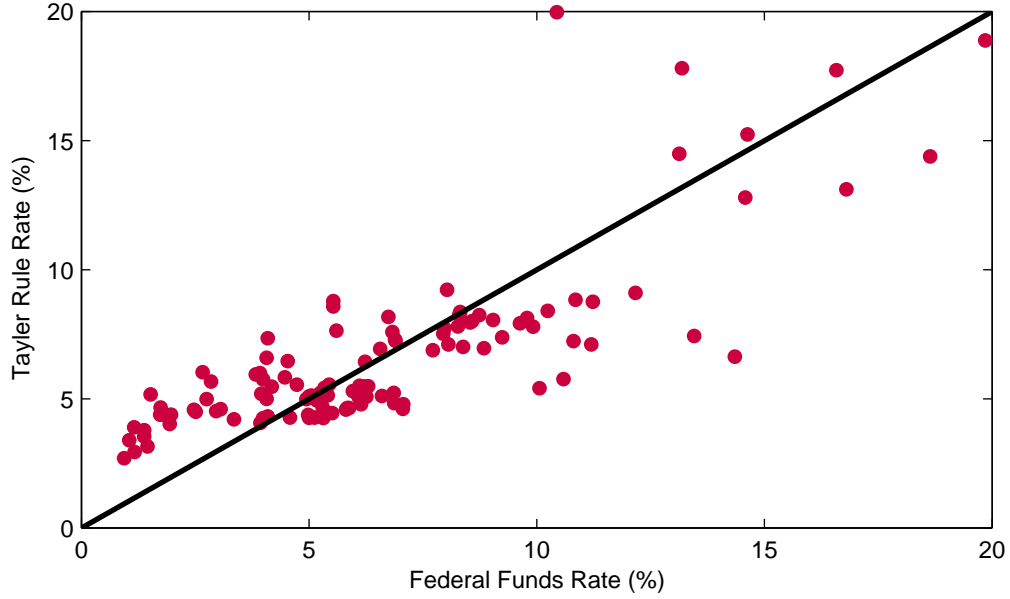
Figure 2: **Federal Funds Rate and Taylor Rule.** This figure plots the observed Federal Funds Rate on a quarterly basis against the the rate prescribed by the simple Taylor rule (22).

FOMC meeting time T. The price of such a contract is given by:

$$P_t = E_t^Q \left[\exp \left(- \int_t^T r_s ds \right) f_T \right] = E_t^Q \left[\exp \left(- \int_t^T r_s ds \right) G(\beta \cdot X_T) \right]. \tag{24}$$

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

$$f(\alpha \cdot X) = 1 \tag{25}$$
$$g(\beta \cdot X) = G(\beta \cdot X_T) \tag{26}$$

3.3 Default Contagion

Next, we consider a model of correlated default that violates the doubly stochastic assumption in that the default event of one firm affects the default probability of another firm beyond its own default intensity. This model is an extension of the primary-secondary framework for counterparty risk in Jarrow and Yu (2001).

Suppose two firms A, B are from the same industry. Firm A is a primary player in the industry, while B is a secondary small firm. Let the default times of A and B be τ_A and τ_B. Under the doubly stochastic assumption, conditional on the information in the driving filtration that determines the

10

intensities λ_A and λ_B, τ_A and τ_B are independent.

$$\Pr(\tau_A > s | \mathcal{F}_t) = E_t \left[\exp\left(-\int_t^s \lambda_A(X_u)\, du\right) \right] \tag{27}$$

$$\Pr(\tau_B > s | \mathcal{F}_t) = E_t \left[\exp\left(-\int_t^s \lambda_B(X_u)\, du\right) \right] \tag{28}$$

However, imagine that the event of the primary firm's default brings major impacts on the rest of the firms in the industry, while the failure of the secondary firm has little effect on others. The impact of A on B can be due to direct business ties that are more important to B than to A, a form of counterparty risk in Jarrow and Yu (2001), or it could be due to A's default changes the perception of risk in other firms, as in Collin-Dufresne et al. (2002). In this case, τ_A remains solely dependent on λ_A, while τ_B depends on λ_B and the status of A.

$$\begin{aligned}
\Pr(\tau_B < s | \mathcal{F}_t) &= \Pr(\tau_B < s, \tau_A > \tau_B) + \Pr(\tau_B < s, \tau_A < \tau_B) \\
&= E_t \left[\int_t^s e^{-\int_t^u \mu(X_v)\,dv} \lambda_B(X_u)\, du \right] \\
&\quad + E_t \left[\int_t^s 1_{\{\tau_A > u, \tau_B > u\}} \lambda_A(X_u) \cdot \Pr(\tau_B < s | \mathcal{F}_{u^-}, \tau_A = u, \tau_B > u)\, du \right]
\end{aligned} \tag{29}$$

where

$$\mu(X) = \lambda_A(X) + \lambda_B(X).$$

Let $\lambda_{B,t}^+$ be the default intensity of B after A defaults. Then, the probability that B defaults before s conditional on A defaulting at time $u < s$ is:

$$\Pr(\tau_B < s | \mathcal{F}_{u^-}, \tau_A = u, \tau_B > u) = 1 - E_u \left[\exp\left(-\int_u^s \lambda_B^+(X_u)\, du\right) \right].$$

Jarrow and Yu (2001) consider the following specification:

$$\lambda_{B,t}^+ = \lambda_{B,t} + c \tag{30}$$

The restriction is that A's default *permanently* increases the default intensity of B by a *constant amount*, which can be relaxed in several ways. For example, we can assume that

$$\lambda_{B,t}^+ = \lambda_{B,t} + c(X_t) \tag{31}$$

for some c which is affine in X. In this case, the second expectation in (29) can be computed in closed form (up to a single integral) using the DPS transform.

Alternatively, we can directly assume that

$$\Pr(\tau_B < s | \mathcal{F}_{u^-}, \tau_A = u, \tau_B > u) = g(\alpha(s - u), \beta(s - u) \cdot X_u). \tag{32}$$

11

where g satisfies the regularity conditions in Theorem 1 plus $\frac{\partial g}{\partial s} > 0$. This specification is attractive when one wants to capture particular empirical patterns of default rates over different horizons (e.g. at the different phases of firm A's bankruptcy process), or the dissipation of the initial impact from A to B over time, which could be infeasible under (31). It could also provide a reduced-form solution to the challenges in modeling the nonlinear dependence of default intensity on state variables found in Duffie et al. (2007), which can become more prominent in a period of market distress following a large firm's default.

One possible choice for g is a logistic function, as used by Campbell et al. (2008):

$$g\left(X,t\right) = \frac{1}{1+\exp\left(-\alpha(t) - \beta(t)\cdot X\right)}, \tag{33}$$

where $\alpha(t)$ and $\beta(t)$ satisfy

$$\alpha'(t) + \beta'(t) \cdot X > 0.$$

The evaluation of (29) given (32) depends on whether X jumps at the default event of firm A, which corresponds to the "no-jump" condition as discussed by Duffie et al. (1996). If X does not jump, the second term in (29) is equal to:

$$\Pr\left(\tau_B < s, \tau_A < \tau_B\right) = V_t = E_t\left[\int_t^s e^{-\int_t^u \mu(X_v)dv}\lambda_A\left(X_u\right)g\left(\alpha\left(s-u\right), \beta\left(s-u\right)\cdot X_u\right)du\right], \tag{34}$$

for which the generalized transform readily applies. If X jumps at τ_A, then a correction term is needed for (34) (see Proposition 1 of Duffie et al. (1996)):

$$\Pr\left(\tau_B < s, \tau_A < \tau_B\right) = V_t - E_t\left[\Delta V_{\tau_A}\right], \tag{35}$$

which again can computed with the generalized transform.

This simple model of contagion can also be used to capture jumps in default intensities and default correlation across firms following major events, with λ_A being the arrival intensity of such events.

3.4 GMM Estimation of Default Models

The need to compute unconditional and conditional moments of nonlinear functions arise in the method of moments and GMM estimations. Consider a class of econometric models given by

$$E_t\left[g\left(\beta X_{t+s}; \theta\right)\right] = 0, \quad s > 0 \tag{36}$$

After conditioning down, we have

$$E\left[g\left(\beta X_t; \theta\right)\right] = 0 \tag{37}$$

The restriction is that the nonlinear moment $g\left(\cdot\right)$ depends on X only through βX. Such a setup has the advantage of making the effects from different elements of X easy to interpret, but does lose the

generality... If X follows the affine jump-diffusion in (), and g satisfies the regularity conditions in Theorem 2.1, then we can compute the above moments using the generalized transform technique. In addition, the framework allows for instruments

$$z_t = f\left(\alpha X_t; \theta\right),$$

such that

$$E\left[g\left(\beta X_t; \theta\right) f\left(\alpha X_t; \theta\right)\right] = 0.$$

Next, we illustrate how to use the generalized transform analysis in a GMM estimaiton of a logit model of default. Suppose the marginal probability of default between t and $t+1$ follows a logistic distribution (see, e.g., Campbell et al 2008)

$$E_t\left[Y_{i,t+1}\right] = \Pr\left(t < \tau_i \le t+1 | X_{i,t}\right) = \frac{1}{1 + \exp\left(-\alpha - \beta X_{i,t}\right)}$$

where τ_i is the time of default for firm i, $Y_{i,t+1} = 1$ if default occurs between t and $t+1$ and 0 otherwise, and $X_{i,t}$ is a vector of explanatory variables which follow the affine diffusion process (). The standard method to estimate the parameters of the logit model is through a logit regression. However, suppose we want to jointly estimate the parameters for the stochastic process X_t and $\{\alpha, \beta\}$ with GMM, then the moments we can use include:

1. the average one-year default probability:

$$E\left[Y_{i,t}\right] = E\left[\frac{1}{1 + \exp\left(-\alpha - \beta X_{i,t}\right)}\right] = \widehat{\mu}_Y = \frac{1}{T} \sum_{t=1}^{T} \frac{\sum_{i=1}^{N_t} Y_{i,t+1}}{N_t}$$

2. the volatility in one-year default probability:

$$std\left[Y_{i,t}\right] = \left[E\left[\left(\frac{1}{1 + \exp\left(-\alpha - \beta X_{i,t}\right)}\right)^2\right] - E\left[Y_{i,t}\right]^2\right]^{\frac{1}{2}} = \widehat{\sigma}_Y = \left(\frac{1}{T} \sum_{t=1}^{T} \left(\frac{\sum_{i=1}^{N_t} Y_{i,t+1}}{N_t} - \widehat{\mu}_Y\right)^2\right)^{\frac{1}{2}}$$

3. additional moments in X

4 Three Concrete Examples

In this section, we provide in-depth analysis of three example applications of the generalized transform analysis.

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 model and illustrate the impact of state-dependent recovery risks on pricing.

We fix a probability space $(\Omega, \mathcal{F}, P)$ and a filtration $\{\mathcal{F}_t : t \geq 0\}$. Following the reduced-form models, the default time is assumed to be a totally inaccessible F-stopping time $\tau : \Omega \to (0, +\infty]$. For simplicity, we assume that τ is doubly-stochastic with intensity λ_t under the risk neutral measure Q.[6] Let the instantaneous riskfree rate be r_t.

The price of a T year defaultable zero-coupon bond with face value of 1 and recovery at default of φ_τ is:

$$
\begin{aligned}
V_0 &= E^Q\left[e^{-\int_0^\tau r_u du}1_{\{\tau \leq T\}}\varphi_\tau\right] + E^Q\left[e^{-\int_0^T r_u du}1_{\{\tau > T\}}\right] \\
&= E^Q\left[\int_0^T e^{-\int_0^t (r_u + \lambda_u)du}\lambda_t \varphi_t dt + e^{-\int_0^T (r_u + \lambda_u)du}\right],
\end{aligned}
\tag{38}
$$

the second equality follows from the doubly-stochastic assumption and certain regularity conditions for r and λ (see Duffie (2005) for details).

Duffie and Singleton (1999) discuss three types of recovery modeling:

1. "recovery of treasury" (RT):
$$
\varphi_t = (1 - L_t)P_t,
\tag{39}
$$
   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):
$$
\varphi_t = (1 - L_t)F,
\tag{40}
$$
   where F is the face value of the bond.

3. "recovery of market value" (RMV):
$$
\varphi_t = (1 - L_t)V_{t-},
\tag{41}
$$
   where V_{t-} is the market value of the security immediately before default.

Duffie and Singleton (1999) show that under the RMV specification and a suitable no-jump condition,[7] 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, the standard results for affine term structure models apply. In contrast, the RT and RFV models are generally not analytically tractable.

---

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

[7]See also Duffie et al. (1996) and Collin-Dufresne et al. (2004) for discussions on the no-jump condition.

While analytically appealing, the RMV assumption also has some limitations. Since λ_t and 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. Moreover, in some cases, e.g. when pricing bonds of different seniorities from the same issuer (e.g. see Figure 1 of Duffie and Singleton 1999), it is more natural to model default intensity λ (same across different bonds) and recovery rates L (depending on seniority). Finally, 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 observed roughly 30 days after the date of default to par value.

Bakshi et al. (2006) study a class of RT and RFV models for which φ is exponential affine in the default intensity. They solve for bond prices using the DPS transform. We show that a wide range of RFV models is also tractable using the generalized transform analysis. The added flexibility in modeling recovery rates allows us to introduce explicit dependence of recovery rates on macro and firm-specific variables.

4.1.1 Model Setup

We directly specify the dynamics of state variables under the risk neutral probability measure. The default intensity λ_t of a firm follow a CIR process,

$$d\lambda_t = \kappa_\lambda(\theta_\lambda - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dB_t^\lambda. \tag{42}$$

The second state variable, Y_t, follows

$$dY_t = \kappa_Y(\theta_Y - Y_t)dt + \sigma_Y\sqrt{\lambda_t}dB_t^Y. \tag{43}$$

The short term interest rate, r_t, is given by

$$r_t = Y_t - \delta\lambda_t. \tag{44}$$

This setup captures the negative correlation between r_t and λ_t as documented by Duffee (1998).

The recovery function φ can depend on the default intensity, short rate, and other macro and firm-specific variables. For example, Altman et al. (2005) document significant negative correlation between aggregate default rates and recovery rates.[8] Chen (2008) shows that macro variables such as GDP growth and riskfree rate are correlated with the aggregate recovery rates and default rates. Zhang (2009) shows that stricter covenants in recessions lead to a negative dependence of recovery rates on lagged macroeconomic conditions. Carey and Gordy (2007) show that firm-level recovery rate increases with the share of bank loans in total debt.

Another important property for the recovery function φ is that it should only take values

---

[8]See Altman et al. (2005) for a review of earlier studies on the relationship between recovery rates and default rates.

from $[0,1]$. One specification for φ that satisfies this requirement and still works with the DPS formulation is $\varphi(X) = e^{\beta \cdot X} 1_{\{\beta \cdot X > 0\}} + 1_{\{\beta \cdot X < 0\}}$. Bakshi et al. (2006) study such a setting. More generally, the cumulative distribution function of any distribution will have this property. CDF functions also have the additional benefit in that they have nice Fourier transform properties.[9] Some of the common choices in empirical studies include:

- 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:

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

The Probit and Cauchy model also have the convenient feature that the integrands have a closed-form Fourier transform.

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

$$\varphi(\lambda) = \frac{a}{1 + b(\lambda - \lambda_0)^2} + c. \tag{45}$$

Re-writing as in the cauchy density:

$$\varphi(\lambda) = \frac{a\pi}{\sqrt{b}} \frac{1}{(\pi b^{-1/2})(1 + (\lambda - \lambda_0)^2/(b^{-1/2})^2)}. \tag{46}$$

Its Fourier transform is:

$$\hat{\varphi}(t) = \frac{a\pi}{\sqrt{b}} e^{\lambda_0 i t - \frac{1}{\sqrt{b}}|t|} \tag{47}$$

We consider two calibrations of $\varphi(\lambda)$ in (45). First, using data from Altman et al. (2005), we calibrate $a = 0.5$, $b = 1415$, $c = 0.25$, and $\lambda_0 = -0.01$. The fitted function is "model I" in Figure 3. 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: the recovery rate is bounded from below at 25%, and it treats the recovery rates in the data the same as risk-neutral recovery rates. Thus, we also study a second calibration, where $a = 0.8$, $b = 800$, $c = 0$, and $\lambda_0 = -0.08$. The fitted function is "model II" in Figure 3, which has similar recovery rates to "model I" when default intensity is low, but has lower recovery than "model I" when default intensity rises.

---

[9]For example, $\hat{f}'(t) = t\hat{f}(t)$ so differentiation and integration is easy on the fourier side by just multiply/divide by t. So if we want to go from PDF to CDF, on fourier side, we just divide by t.
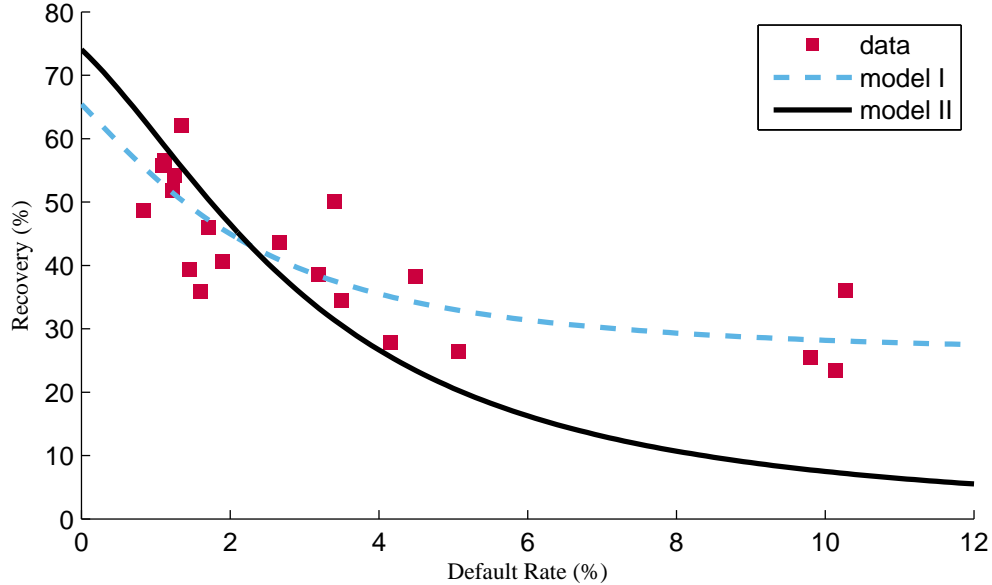
Figure 3: **A Cauchy Model of Aggregate Recovery Rates.** This figure plots the aggregate recovery rates and default rates from Altman et al. (2005). The blue line fits a Cauchy curve through the data. The functional form of the curve is given by equation (45).

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 affect pricing.

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

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

which can be easily mapped into the generalized transform (6),

$$f \left(\alpha \cdot X \right) = \iota_2 \cdot X \tag{48}$$

$$g \left(\beta \cdot X \right) = \frac{a}{1 + b(\iota_2 \cdot X - \lambda_0)^2} + c \tag{49}$$

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 doubly-stochastic assumption or no-jump conditions.[10] Thus, it still works in models with flight-to-quality, default contagion, systematic jump risk, or other features that violate the no-jump condition.

---

[10]We can either explicitly make the correction for jumps as in Duffie et al. (1996), or use the change-of-measure method in Collin-Dufresne et al. (2004).

Table 1: MODEL CALIBRATION

κ_λ	θ_λ	σ_λ	κ_Y	θ_Y	σ_Y	δ
0.1	0.06	0.01	0.02	0.12	0.015	0.1

4.1.2 Results

We use the dynamics of default intensity λ and riskfree rate r implied by (42-44) and the recovery model (45) to price 10-year defaultable zero-coupon bonds. The parameter values for the λ and Y process are reported in Table 1.

The results of the model are reported in Figures 4. Panel A of Figures 4 plots the credit spreads (the yield spreads between the defaultable bond and default-free bond with same maturity) as functions of the default intensity. We report the results for (1) the stochastic recovery model I; (2) constant recovery with φ set to the unconditional mean recovery rate implied by model I, which is 31%; (3) constant recovery with $\varphi = 25\%$, a popular assumption for the recovery rate in both academic analysis and industry practice (see, e.g. Pan and Singleton 2008).

Panel B reports the fraction of the spreads due to recovery risk, defined as

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

which can be interpreted as a measure of the relative pricing errors due to ignoring the state-dependence of recovery rates.

Naturally, credit spreads increase with default intensity in all cases. When current default intensity is low, the stochastic recovery model implies higher-than-average recovery rates. As a result, constant recovery model generates higher credit spreads. As default intensity rises, the spreads from the stochastic recovery model become higher, and exceeds the spreads from the constant recovery model with φ set to model-implied average recovery rate. Since the lower bound for the stochastic recovery model model is 25%, the spreads of the stochastic recovery model is always below that of the constant recovery model with $\varphi = 25\%$.

In terms of the relative pricing errors, relative to stochastic recovery model I, the main concern of the constant recovery assumption is underpricing, i.e. generating credit spreads that are too high. The pricing errors can be large. When default intensity is low ($< 2\%$), the pricing errors are close to 40% for the assumption of 25% recovery rate.

In Panel C and D of Figures 4, we compare the pricing results of the stochastic recovery model II with the constant recovery models. Again, when default intensity is low, credit spreads are lower for the stochastic recovery model, because the conditional recovery rates are higher than under constant recovery assumption. As default intensity increases, the low recovery rate at times of high default probability starts to have more and more significant effects on the bond prices. The spreads rise more rapidly in model II than in model I (Panel A) due to the more rapid decline in recovery rates, and they exceed the spreads from the constant recovery models quickly.
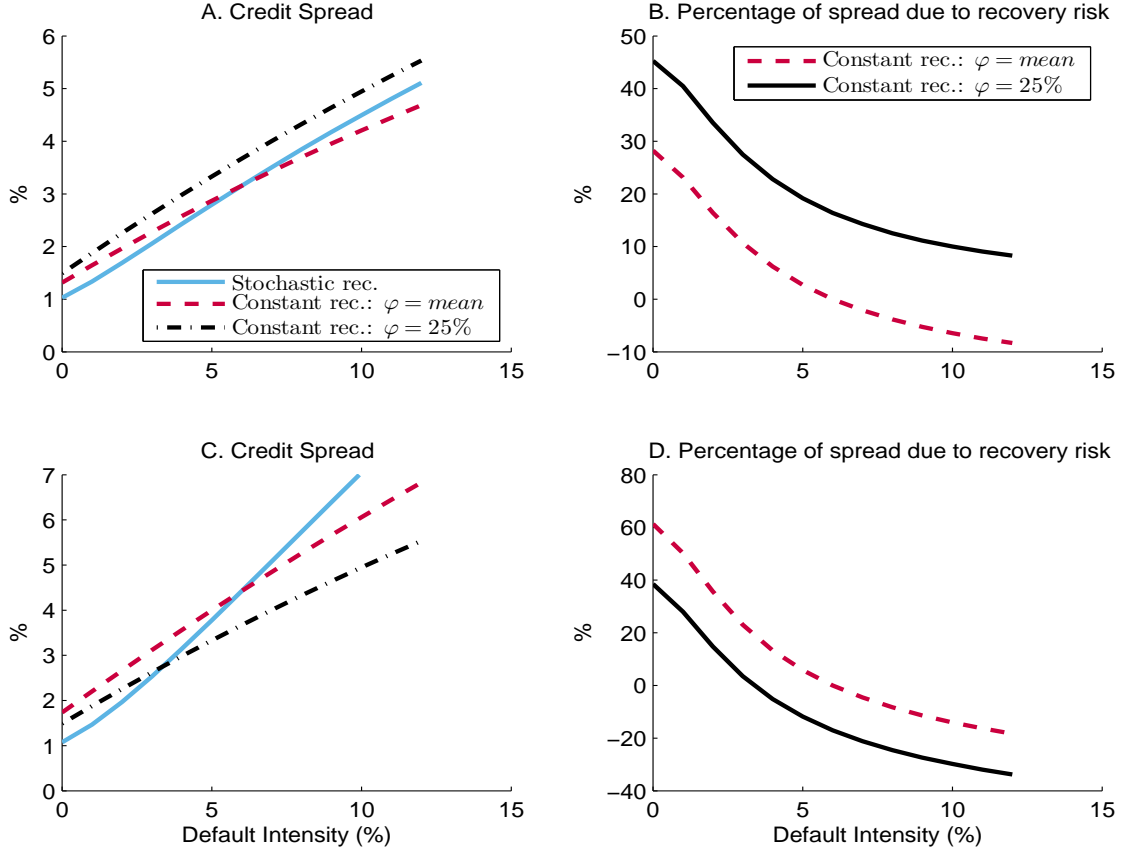
Figure 4: **Credit spreads for 10-year bonds with constant recovery and Cauchy recovery.** This figure plots the credit spreads for 10-year zero-coupon defaultable bonds. The stochastic recovery model in Panel A and B are based on recovery model I, Panel C and D based on recovery model II. For constant recovery, "$\varphi = mean$" sets the recovery rate to the unconditional mean of the stochastic recovery model, which is 31% for model I, 16% for model II.

The relative pricing errors in Panel D are quite large either when default intensity is low or high. For example, under the 25% constant recovery rate assumption, the spreads are 30%+ higher than the stochastic recovery model for $\lambda < 1\%$, and 18%+ lower than the stochastic recovery model for $\lambda > 6\%$. Moreover, simply adjusting the constant recovery rate does not solve the problem since it will either exacerbate the underpricing for low λ or overpricing for high λ. These results suggest that the negative correlation between default intensity and recovery rate can have important impact for pricing high grade bonds as well as high yield bonds.

Finally, we use the Monte Carlo method to check the accuracy of our solutions. The results are reported in Figure 5. The yields computed from the Monte Carlo method (using 50,000 simulations) are consistent with the spreads computed using the generalized transform. The remaining differences are due to the numerical errors of the simulations.
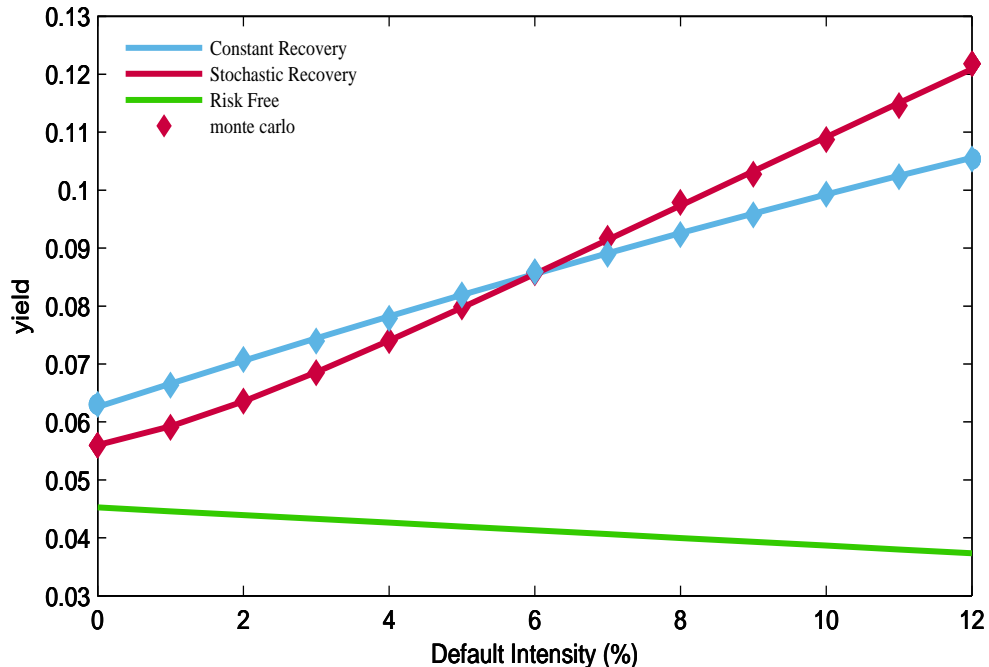
Figure 5: **Zero coupon yields from generalized transform vs. Monte Carlo method.**This figure plots zero coupon yields for 10-year bonds with stochastic recovery and constant recovery, compared to prices obtained using Monte Carlo method.

4.2 Two Non-IID trees

In this section we illustrate how to use the generalized transform to compute the equilibrium of asset pricing models. Cochrane et al. (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 in the case of log utility and i.i.d. trees. Martin (2008) extends the analysis to multiple trees and power utility, but also assume that dividend growth of each tree is i.i.d. We show that the model can be extended to allow for mean reversion, conditional heteroscedasticity, and jumps with time-varying intensity in dividend growth.

4.2.1 The Model

Following Cochrane et al. (2008) and Martin (2008), we consider an endowment economy with two stocks (trees). The infinitely-lived representative investor has CRRA utility:

$$U_t = E_t \left[\int_0^\infty e^{-\rho u} \frac{C_{t+u}^{1-\gamma} - 1}{1 - \gamma} du \right], \tag{50}$$

where we focus on the case $\gamma > 1$.

There are two stocks with dividend streams $D_{1,t}dt$ and $D_{2,t}dt$. Our model deviates from Cochrane et al. (2008) and Martin (2008) in that the dividend growth is *non-i.i.d.* The log dividends $d_{1,t} = \log D_{1,t}$ and $d_{2,t} = \log D_{2,t}$ follow the processes:

$$dd_{i,t} = g_{i,t}dt + \sigma_{d,i}dZ_{i,t}^d \tag{51}$$

$$dg_{i,t} = \kappa_i\left(\overline{g}_i - g_{i,t}\right)dt + \sigma_{g,i}dZ_{i,t}^g \tag{52}$$

where $g_{i,t}$ is the expected growth rate for $d_{i,t}$, which follows an Ornstein-Uhlenbeck process with long term mean $\overline{g}_i$. For simplicity, we assume that all the Brownian motions $Z_{i,t}^d$ and $Z_{i,t}^g$ are uncorrelated with each other. When $\gamma = 1$ and $g_{i,t} \equiv \overline{g}_i$, we recover the model of Cochrane et al. (2008). When $\gamma > 1$ and $g_{i,t} \equiv \overline{g}_i$, we recover the two-tree model of Martin (2008).

In equilibrium, aggregate consumption $C_t = D_{1,t} + D_{2,t}$. Define the dividend share for the first stock as $S_t = D_{1t}/(D_{1t} + D_{2t})$. As in the Lucas economy with a single asset, the instantaneous riskfree rate in this economy is determined by the rate of time preference, the expected growth rate of consumption, and precautionary savings driven by consumption volatility:

$$r_{f,t} = \rho + \gamma\left(S_t(g_{1,t} + \sigma_{d,1}^2/2) + (1 - S_t)(g_{2,t} + \sigma_{d,2}^2/2)\right) - \frac{\gamma(\gamma + 1)}{2}\left(S_t^2\sigma_{d,1}^2 + (1 - S_t)^2\sigma_{d,2}^2\right). \tag{53}$$

Under the standard regularity conditions, the price of stock i ($i = 1, 2$), $P_{i,t}$, is given by:

$$\begin{aligned}
P_{i,t} &= E_t\left[\int_0^\infty e^{-\rho u}\frac{u_c\left(C_{t+u}\right)}{u_c\left(C_t\right)}D_{i,t+u}du\right] \\
&= E_t\left[\int_0^\infty e^{-\rho u}\frac{(D_{1,t} + D_{2,t})^\gamma D_{i,t+u}}{(D_{1,t+u} + D_{2,t+u})^\gamma}du\right] \\
&= (D_{1,t} + D_{2,t})^\gamma \int_0^\infty e^{-\rho u}E_t\left[\frac{D_{i,t+u}}{(D_{1,t+u} + D_{2,t+u})^\gamma}\right]du. \tag{54}
\end{aligned}$$

To map the expectation in (54) to the generalized transform formula, we first define the state variable $X_t = (d_{1t}, \mu_{1t}, d_{2t}, \mu_{2t})$. The model (51)-(52) implies that X follows an affine diffusion process,

$$dX_t = (K_0 + K_1 X_t)\,dt + \sqrt{H_0}\,dZ_t, \tag{55}$$

where

$$K_0 = \begin{pmatrix} 0 \\ \kappa_1\overline{g}_1 \\ 0 \\ \kappa_2\overline{g}_2 \end{pmatrix}, \quad K_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -\kappa_1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -\kappa_2 \end{pmatrix}, \quad H_0 = \begin{pmatrix} \sigma_{d,1}^2 & 0 & 0 & 0 \\ 0 & \sigma_{g,1}^2 & 0 & 0 \\ 0 & 0 & \sigma_{d,2}^2 & 0 \\ 0 & 0 & 0 & \sigma_{g,2}^2 \end{pmatrix}$$

$$Z_t = \left[Z_{1,t}^d\; Z_{1,t}^g\; Z_{2,t}^d\; Z_{2,t}^g\right]'$$

Next, from (54),

$$E_t \left[\frac{D_{1,s}}{(D_{1,s} + D_{2,s})^\gamma} \right] = E_t \left[\frac{e^{(1-\gamma/2)d_{1,s} - \gamma/2 d_{2,s}}}{\left(2 \cosh \frac{d_{1,s} - d_{2,s}}{2} \right)^\gamma} \right] = E_t \left[f\left(\alpha \cdot X_s \right) g\left(\beta \cdot X_s \right) \right], \tag{56}$$

where

$$f(x) = e^x \tag{57}$$

$$g(x) = \frac{1}{(2 \cosh(x))^\gamma} \tag{58}$$

and

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

Up to this point, our procedure is the same as in Martin (2008). When the increments of X are $i.i.d.$, the conditional characteristic function for X is known explicitly, which Martin uses to compute (56) following a Fourier transform for g. We generalizes his method to the case where the increments of X are $non\text{-}i.i.d.$ by exploiting the properties of the conditional characteristic function for general affine processes.

This model can be further extended in several dimensions. First, while here we only consider time-varying expected growth rates, X can be any affine process, which allows us to introduce stochastic volatility and jumps with time-varying intensity in consumption growth. One attractive feature of our approach is that these new elements do not increase the dimension of the Fourier transform, which maintains the tractability of the model. 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 et al. (2007), where the two trees are interpreted as nonhousing consumption and housing services. We can also impose cointegration between the two dividend process and allow for stochastic volatility, as do Piazzesi et al. (2007). Third, it is also convenient to add preference shocks that are $pl\text{-}linear$ in the state variable. Finally, the model can allow for multiple trees using the multi-dimensional version of the generalized transform in Proposition 2.

4.2.2 Price-Dividend Ratios and Expected Excess Returns

To illustrate the quantitative implications of the model, we choose the following parameters. For preferences, $\rho = 0.1$ and $\gamma = 10$. For the two dividend processes, we assume that the first stock initially has smaller dividend than the second stock, $D_{1,t} = 1/5 D_{2,t}$, but has higher growth rate in the long run, $\bar{g}_1 = 0.03 > \bar{g}_2 = 0.01$. The other parameters are the same for the two stocks: the speed of mean reversion for the growth rate $\kappa_1 = \kappa_2 = 0.3$, volatility of dividend $\sigma_1^d = \sigma_2^d = 0.1$, and volatility of dividend growth $\sigma_1^g = \sigma_2^g = 0.01$. A possible interpretation for this parametrization is that a young, fast-growing industry is slowly taking over the economy from an established, slow-growing industry.
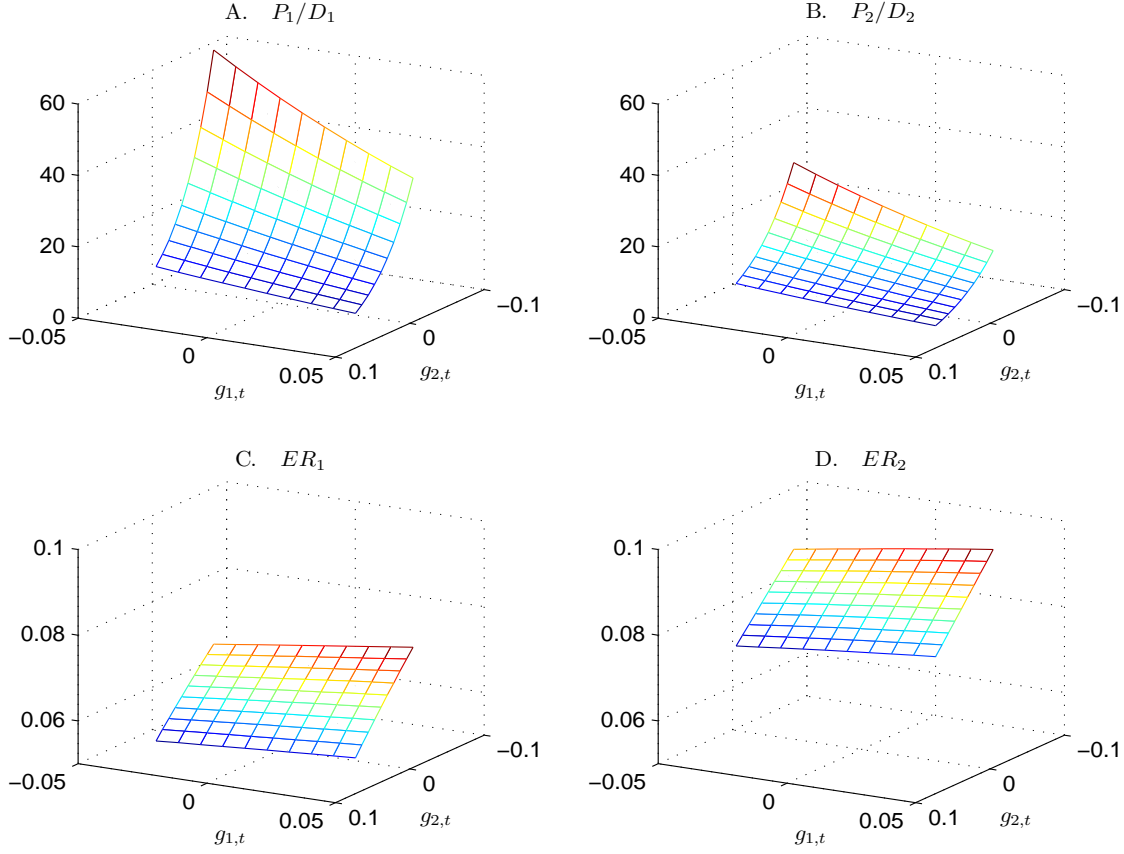
Figure 6: **Price-dividend ratios and expected excess returns.** This figures plots the price-dividend ratios and expected excess returns of the two stocks $(D_{1t}/D_{2t} = 1/5)$ as a function of their own dividend growth rate and the other stock's growth rate.

The top two panels of Figure 6 plot the price-dividend ratios for the two stocks against the conditional expected growth rates of the two stocks. The price-dividend ratio the first stock is higher than the second, and it is decreasing in the expected growth rate of both stocks, $g_{1,t}$ and $g_{2,t}$, although the decline is faster with $g_{2,t}$. A rise in the expected growth rate has two effects on the price-dividend ratio of the stock. There is a *cash-flow effect*: higher expected growth rate implies higher future cash flows, which tends to increase the price-dividend ratio. There is also a *discount rate effect*: a rise in the expected growth rate will increase the discount rate for stocks due to the "substitution effects": when the expected future consumption is high, consumers want to borrow to increase consumption today, which raises the interest rate (see equation (53)). With the CRRA utility and the given parameters, the discount rate effect dominates, causing price-dividend ratio to decrease with the growth rates.

Due to the large size of the second stock, an increase in $g_{2,t}$ has a strong effect on the aggregate discount rate, which causes the price of the first stock to drop. This effect is stronger when $g_{1,t}$ is small. In contrast, changes in $g_{1,t}$ have a small effect on the discount rates, but it has the additional

cash flow effect on the small stock. As the share of the first stock gets smaller in this economy, the cash flow effect eventually dominates, resulting in the price-dividend ratio $P_{1,t}/D_{1,t}$ increasing in $g_{1,t}$. As for the large stock, when $g_{2,t}$ is large, changes in $g_{1,t}$ have almost no effect on the price-dividend ratio $P_{2,t}/D_{2,t}$. But when $g_{2,t}$ is small, the effect of $g_{1,t}$ on the price of the larger stock can become sizable.

In order to analyze the expected excess returns, we consider a stock as a portfolio of zero-coupon equities. The risk premium of the stock is then the value-weighted average of the risk premium for these zero-coupon equities, which are easy to compute. In general, the instantaneous expected excess return for any asset is determined by its exposure to the primitive risk sources, $X_t = (d_{1t}, g_{1t}, d_{2t}, g_{2t})$, and the risk premia that these risks demand through their covariance with the pricing kernel. In this case, the pricing kernel, M_t, takes the form $M_t = e^{-\rho t}(D_{1,t} + D_{2,t})^\gamma$. By Îto's Lemma, the expected excess returns for any asset with state-dependent price $P_t = P(X_t, t)$ are then given by

$$
\begin{aligned}
ER_t &= \frac{1}{P}\frac{E^P[dP]}{dt} - r_t \\
&= \frac{1}{P}\nabla_X P(X_t) \cdot H_0 \nabla_X M(X,t)/M(X,t) \\
&= \frac{1}{P}\frac{\partial P}{\partial d_1}\gamma S_t \sigma_{d,1}^2 + \frac{1}{P}\frac{\partial P}{\partial d_2}\gamma(1 - S_t)\sigma_{d,2}^2,
\end{aligned}
\tag{59}
$$

where $M(X,t) = e^{-\rho t}(e^{X_1} + e^{X_3})^\gamma$ is the stochastic discount factor and $S_t = D_{1t}/(D_{1t} + D_{2t})$ is the dividend share. Since the innovations in dividends and growth rates are uncorrelated, there is no premium for growth risk in this model. Notice that the risk premium of a stock does not necessarily go to zero when its share approaches 0. Even though the stock's own dividend shocks become uncorrelated with the pricing kernel as its share drops to 0, the stock can still be exposed to discount rate rate risks. For example, a shock to the second stock's dividend will change the share and the diversification in the economy, which affect the discount rates, and in turn, the price of the first stock.

Since the price of zero coupon equity is available in closed form through the transform analysis, its gradient can also be computed in closed form. Then we can compute the risk premium for zero coupon equities at different maturities using (59), and compute the risk premium for the stock as a weighted average of the premia for zero coupon equities. The results are plotted in the bottom panels of Figure 6. Because the second stock is larger and thus more correlated with the pricing kernel, it demands higher risk premium on average. Interestingly, the conditional risk premia of both stocks are decreasing in the growth rate of the larger stock $g_{2,t}$, but increasing in the growth rate of the smaller stock $g_{1,t}$.

Equation (59) implies that the expected growth rates of dividend only affect the risk premia by changing the sensitivity of the log stock price to dividend shocks. A positive shock to $d_{1,t}$ increases expected future dividends for the first stock, which tends to raise its stock price, and moves the share S_t closer to 0.5, which increases the interest rate and lowers the stock price. A positive shock

to $d_{2,t}$ has no impact on future dividends of the first stock, but moves the share S_t closer to 0, which decreases the interest rate and raises the stock prices. The net risk premium of the smaller stock is more affected by the exposure to the dividend shock of the larger stock, resulting in positive risk premium. A higher $g_{1,t}$ will amplify the effects of both dividend shocks on the interest rates, but since the second dividend shock bears a higher premium, the net risk premium increases in $g_{1,t}$. A higher $g_{2,t}$ reduces the effects of both dividend shocks on the interest rates and prices, which causes the risk premium to drop.

4.3 Differences of Opinions

One is often interested in economies where there is heterogeneity amongst agents in terms of beliefs, or equivalently of preferences. These differences generate trade and affect asset prices in equilibrium. In studying such economies, aggregation often lead 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.

Suppose that there are heterogeneous agents ($i = 1, 2$ for simplicity) which possess heterogeneous beliefs. There is a state variable X_t which Agent 1 believes follows an affine jump diffusion:

$$dX_t = \mu_t^1 dt + \sigma_t^1 dB_t^1 + dZ_t^1 \tag{60}$$

where $\mu_t^1 = K_0^1 + K_1^1 X_t$, $\sigma_t^1 (\sigma_t^1)^\top = H_0^1 + H_1^1 \cdot X_t$ and jumps are believed to arrive with intensity $\lambda_t^1 = \lambda_0^1 + \lambda_1^1 \cdot X_t$ and have distribution ν^1 (with moment generating function ϕ^1). As elaborated in the examples below, the variable X_t is meant to encompass all uncertainty in the economy including any time-variation in the heterogeneity of beliefs. For simplicity, we suppose that Agent 1's beliefs are correct.

Agent 2 has heterogeneous beliefs which we shall suppose are equivalent. A broad class[11] of such equivalent beliefs can be characterized as follows. There exists some vector a such that Agent 2 believes X follows an affine jump diffiusion satisfying

1. $\mu^2 = \mu^1 + \sigma^1 (\sigma^1)^\top a$

2. $\sigma^2 = \sigma^1$

3. $d\nu^2 / d\nu^1 (Z) = e^{a \cdot Z} / E_{\nu^1}[e^{a \cdot Z}]$ or $\phi^2(c) = \phi^1(c + a) / \phi_1(a)$

4. $\lambda^2 = \lambda^1 \times E_{\nu^1}[e^{a \cdot Z}]$

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.

---

[11]More generally, we could consider beliefs of the form $e^{h(x_t) - \int_0^t \mathcal{D}^1 h(x_s) ds}$. Provided the integral term remains tractable, the same analysis applies.

This structure implies that the two beliefs define equivalent probability measures which may be related through the Radon-Nikodym derivative dP^2/dP^1

$$\eta_t = E_t[\frac{dP^2}{dP^1}].$$
$$= e^{a \cdot X_t - \int_0^t a \cdot \mu_s^1 + \frac{1}{2}\|\sigma_s^1 a\|^2 + \lambda_s^1(\phi_\nu(a)-1)ds} \tag{61}$$

The variable η_t expresses Agents 2's differences in opinion in that when η_t is high, Agent 2 believes an event is more likely than Agent 1 believes. To distinguish from the risk neutral measure, we refer to the η_t defining the difference in beliefs the db-density process.

Suppose now that the agents have time separable utility:

$$U_i(c) = E_0[\int_0^\infty u(c_t, t)dt] \tag{62}$$

Suppose also that (1) markets are complete, (2) log of aggregate consumption, c_t, is linear in X_t (say, $c_t = c \cdot X_t$), (3) agents are endowed with some fixed fraction $(\theta_1, \theta_2 = 1 - \theta_1)$ of aggregate consumption. As in Dumas et al. (2009), hereafter DKU, we can solve for equilibrium prices. Let ξ_t denote the density process $E_t[dQ/dP^1]$, where Q is the risk-neutral measure. As in ?, we can set the lifetime budget constraint and equate state prices to marginal utilities to solve

$$U_t'(c_t^1) = \lambda^1 \xi_t , \tag{63}$$
$$U_t'(c_t^2) = \eta_t \lambda^2 \xi_t . \tag{64}$$

Market clearing then implies

$$c_t = U_t^{-1}(\lambda_1 \xi_t) + U_t^{-1}(\lambda_2 \eta_t \xi_t) \tag{65}$$

where c_t^i is Agent i's equilibrium consumption at time t, λ^i is the shadow price of an additional util for Agent i. This gives a nonlinear equation for η_t, so that in general $\xi_t = h(c_t, \eta_t)$, where h solves the associated inverse problem.

With the additional assumption that $U_t(c) = e^{-\rho t}c^{1-\gamma}/(1-\gamma)$, this simplifies to

$$U_t'(c) = e^{-\rho t}c^{-\gamma} \tag{66}$$
$$c_t^1 = (e^{\rho t}\lambda_1 \xi_t)^{-1/\gamma} \tag{67}$$
$$c_t^2 = (e^{\rho t}\lambda_1 \xi_t/\eta_t)^{-1/\gamma} \tag{68}$$
$$\xi_t = e^{-\rho t}\left[\left(\frac{\eta_t}{\lambda_1}\right)^{1/\gamma} + \left(\frac{1}{\lambda_2}\right)^{1/\gamma}\right]^\gamma c_t^{-\gamma} \tag{69}$$

Using $g(x) = \left[\left(\frac{x}{\lambda_1}\right)^{1/\gamma} + \left(\frac{1}{\lambda_2}\right)^{1/\gamma}\right]^{\gamma}$ and writing $c_t = e^{c \cdot X_t}$ we have

$$\xi_t = e^{-\rho t} g(a \cdot X_t) e^{c \cdot X_t} \tag{70}$$

Thus with the risk neutral density in this form we may price any asset with pl-linear payoff, such as bonds and dividend claims, using Theorem 1.[12]

We now provide a number of examples showing that this framework 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.

1. **Disagreement about Stochastic Growth Rates.** This is the example studied in DKU. We now show how this framework corresponds to that of DKU. As we elaborate on later, our main results substantially simplify the calculations for the most general model that they consider. In their model, there is a single dividend tree with time-varying growth rate, but heterogeneous agents with differing beliefs regarding the growth rate of the tree. Their model can be mapped into our setting by using the state variable $X_t = \langle \log \delta_t, \log \eta_t, \hat{f}_t^B, \hat{g}_t, \hat{g}_t^2 \rangle$, where δ_t represents dividends, η_t denotes what they call sentiment, $\hat{f}_t^B$ represents beliefs about the growth rate of dividends for group B of agents and $\hat{g}_t$ represents differences in opinions about the growth rate of dividends among the heterogenous agents. The dynamics of X_t are given by the stochastic differential equation:

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

where

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

It is easy to see that the local conditional variance of X_t, $\Sigma_t \Sigma_t^\top$, is affine in X_t so this represents an affine process.[13]

---

[12]The 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.

[13]DKU 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.

The variable η_t who they refer to as sentiment also gives the density process: $\eta_t = E_t[dP^A/dP^B]$. It is immediate that η_t takes the form of (61) 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). In the main text of their paper, they consider the case where $b \in \mathbb{N}$ so that the nonlinearities in g can simply be expanded into a sum of exponential terms thereby reducing to log-linear functionals. In this case, prices can be computed by well-known methods. For more general cases, they derive expressions in their appendix. We now show how our method simplifies the dimensionality of their analysis for the nonlinear case.

The formula they present in (A58-A61) is (essentially)

$$E_0[e^{\alpha \cdot X_t} g(\beta \cdot X_t)] = \frac{1}{2\pi} \int_{b \in \mathbb{R}} g(b) \int_{s \in \mathbb{R}} e^{ibs} E_0[e^{(\alpha - is\beta) \cdot X_t}] ds \ db. \tag{71}$$

This gives an alternative expression, a *double* integral, for the expectations that we compute using a *single* integral. Equation (71) can be derived by a (more-or-less) direct application of the results in DPS. However, as noted, this formula increases the dimensionality of the problem. Theorem 1 shows that such nonlinearities need not increase the dimensionality of the problem.[14]

2. **Disagreement about Volatility Risk.** Say that dividends have stochastic volatility:

$$X_t = [c_t, V_t]^\top \tag{72}$$

$$dX_t = \begin{bmatrix} \bar{g} \\ -\kappa_V V_t \end{bmatrix} dt + \sqrt{\begin{bmatrix} \sigma_d & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \sigma_{cV} V_t & 0 \\ 0 & \sigma_{VV} V_t \end{bmatrix}} dB_t^1 \tag{73}$$

Here σ_d^2 is the lowest conditional variance of log dividends, while V_t represents the degree to which volatility is above the lowest level.

Agent 2 disagrees about the dynamics of volatility. According to their beliefs:

$$dX_t = \begin{bmatrix} \bar{g} \\ -(\kappa_V - \gamma) V_t \end{bmatrix} dt + \sqrt{\begin{bmatrix} \sigma_d & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \sigma_{cV} V_t & 0 \\ 0 & \sigma_{VV} V_t \end{bmatrix}} dB_t^2 \tag{74}$$

Thus when $\gamma > 0$, these agents believe that volatility mean reverts more slowly. Using $a = \langle 0, \gamma/\sigma_{VV}^2 \rangle$ gives the approach change of measure from Equation (61). In this case, $\langle X_t, I_t \rangle$ follows a 3-dimensional affine process.

3. **Disagreement about Momentum.** Suppose that there is stochastic growth in consump-

---

[14]Indeed for the generalization $g(\beta_1 \cdot X_t, \beta_2 \cdot X_t)$ the trade-off is a somewhat tractable 2-dimensional integral with our method versus a highly intractable 4-dimensional integral by using a direct extension of the DPS method.

tion.

$$\begin{bmatrix} c_t \\ g_t \\ e_t \end{bmatrix} = \begin{bmatrix} g_t \\ \kappa(\bar{g} - g_t) \\ g_t - re_t \end{bmatrix} + \begin{bmatrix} \sigma_c & 0 \\ 0 & \sigma_g \\ 0 & 0 \end{bmatrix} dB_t^1 \tag{75}$$

In this case, e_t is an exponential weighted moving average of past growth:

$$e_t = \int_{-\infty}^{t} e^{-rs} g_{t-s} ds \tag{76}$$

Suppose now that Agent 2 believes

$$\begin{bmatrix} c_t \\ g_t \\ e_t \end{bmatrix} = \begin{bmatrix} g_t + ce_t \\ \kappa(\bar{g}_2 - g_t) \\ g_t - re_t \end{bmatrix} + \begin{bmatrix} \sigma_c & 0 \\ 0 & \sigma_g \\ 0 & 0 \end{bmatrix} dB_t^2 \tag{77}$$

In this case, Agent 2 believes that growth is due to two components (1) a mean-reverting component, g_t and (2) a counteracting momentum component through e_t. 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 2 will believe that past deviations from the steady state lead to larger future deviations from the state steady. In this way we can very Agent 2 as possessing a conservatism or "law of small numbers" bias.

In this case, the difference in beliefs cannot be directly expressed as in (61) directly. However, by considering an augmented state variable we can return to the form of (61). To see this, let $\hat{X}_t = \langle X_t, \text{vech}(X_t) \rangle$. It is easily verified that $\hat{X}_t$ is an affine process and that

$$d(c_t e_t) = \mu_t^{ce} dt + [\sigma_c e_t, \ 0] dB_t^1 \tag{78}$$

Thus a difference in belief regarding $\hat{X}$ which satisfies (61) so that the coordinate of a corresponding to $c_t e_t$ non-zero with the rest of the entries zero will satisfy (75) and (77). Similarly, the integral term in (61) has affine drift in $\hat{X}_t$. Such techniques are common in the term structure literature with respect to quadratic term structure models. The procedure generalizes to accommodate such an "essentially affine" difference of beliefs (see Duffee (2002)).

4. **Disagreement about Higher order Beliefs.** Within this setting, we see that heterogeneous beliefs affect asset prices. Similarly, 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 $X_t = [c_t, \hat{g}_t^1, \hat{g}_t^2, \hat{g}_t^{12}, \hat{g}_t^{21}]$. Let us suppose that X_t follows a Gaussian process under both agents beliefs. Consistency requires that the drift of $\hat{g}_t^2$ under Agent 1's beliefs be $\hat{g}_t^{21}$. Similarly, we have a consistency requirement for Agent 2. For example, we may suppose that Agent 1's beliefs are such that

$$dX_t = \begin{bmatrix} \hat{g}_t^1 \\ \kappa_1(\theta - \hat{g}_1) \\ \hat{g}_t^{12} \\ \kappa_{11}(\hat{g}^2 - \hat{g}^{12}) \\ \kappa_{12}(\hat{g}^1 - \hat{g}^{21}) \end{bmatrix} + \Sigma dB_t^1 \qquad (79)$$

Here, the fourth and fifth components of the drift say that Agent 1 believes that the higher order beliefs (both his beliefs about Agent 2 and Agent 2's beliefs about him) are correct in the long run, but may have short run deviations.

$$dX_t = \begin{bmatrix} \hat{g}_t^2 \\ \hat{g}_t^{21} \\ \kappa_2(\theta_2 - \hat{g}_2) \\ \kappa_{21}(\hat{g}^2 - \hat{g}^{12}) \\ \kappa_{22}(\hat{g}^1 - \hat{g}^{21}) \end{bmatrix} + \Sigma dB_t^2 \qquad (80)$$

Again, Agent 2 believes that the higher order beliefs converge to the correct beliefs in the long run, but perhaps at a slower rate than Agent 1 believes.

As in the Disagreement About Momentum example, we must extend the state variable to include the cross-product of the state-variables in order to fit into the form of (61).

5. **Disagreement about Disasters.** Suppose that log consumption, c_t, has constant growth with IID innovations with time-varying probability, λ_t, of rare disaster. Let $X_t = [c_t, \lambda_t]$

$$dX_t = \begin{bmatrix} \bar{g} \\ -\kappa_\lambda \lambda_t \end{bmatrix} dt + \begin{bmatrix} \sigma_c & 0 \\ 0 & \sigma_\lambda \sqrt{\lambda} \end{bmatrix} dB_t^1 + dJ_t^1 \qquad (81)$$

where J_t are jumps in c_t which occur with intensity $\lambda_0 + \lambda_t$ and distribution ν. Suppose that Agent 2's beliefs are specifed by the db-density of form (61) with $a = \langle c, 0 \rangle$. Agent 2's beliefs
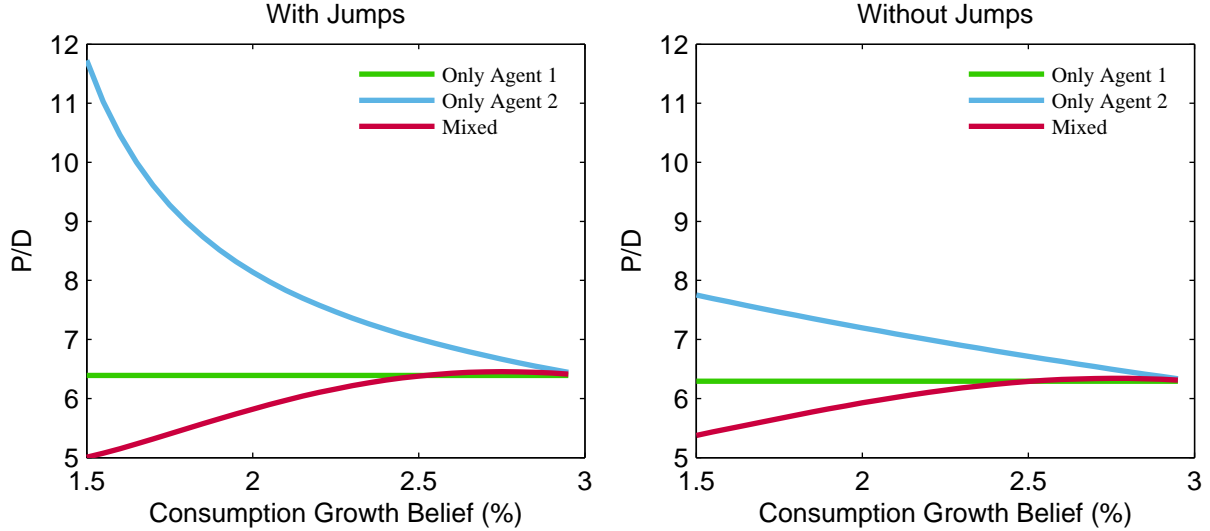
Figure 7: **Price-Dividend Ratios with Heterogeneity in Beliefs.**These figures plots the Price-Dividend ratio in an economy where agents hold different beliefs regarding growth rates and disaster probabilities.

will follow from

$$dX_t = \begin{bmatrix} \bar{g} + \sigma_c^2 a \\ -\kappa_\lambda \lambda_t \end{bmatrix} dt + \begin{bmatrix} \sigma_c & 0 \\ 0 & \sigma_\lambda \sqrt{\lambda} \end{bmatrix} dB_t^2 + dJ_t^2 \tag{82}$$

where jumps arrive with intensity $\lambda_t^2 = E_{\nu^1}[e^{a \cdot Z}](\lambda_0 + \lambda_t)$ and have distribution ν^2 with Radon-Nikodym derivative $d\nu^2/d\nu^1(Z) = e^{a \cdot Z}/E_{\nu^1}[e^{a \cdot Z}]$.

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

For illustration, we consider the disagreement about disasters in Example 5. The model is calibrated as follows:

- $\bar{g} = 3\%, \sigma_c = 2\%, \kappa_\lambda = .1, \theta_\lambda = 1.5\%, \sigma_\lambda^\infty = 1\%$, (the volatility of the stationary distribution of λ), $\gamma = 2$. Jumps are always a 7.5% drop in consumption.

- Agent 2's beliefs are generated with $a = \langle a_1, 0 \rangle$ where a_1 is chosen so that the agent is more pessimistic (both with regard to growth rates and the likelihood of disasters). The believed growth rate is g_2 when $a = (g_2 - \bar{g})/\sigma_c^2$.

- Each agent has fixed endowment of $\theta_1 = \theta_2 = \frac{1}{2}$. This is achieved by adjusting the Lagrange multipliers for the shadow utility of an additional unit of consumption.

Note that this difference in beliefs implies the difference in jump intensity to vary by the multiplicative factor $e^{(g_2 - g_1)/\sigma_c^2 j}$ where j is the jump size. For example, when $g_1 = 2.4\%$, this

factor corresponds to 3.08, meaning the second agent believes that disasters are three times more likely than the optimistic agent.

Figure 7 plots the price-dividend ratio of the asset that pays the consumption claim as dividend. The wealth fraction is maintained at $\theta_1 = \frac{1}{2}$. For small disagreement, the heterogeneous economy lies between the single agent economies. For extreme disagreement, the price becomes depressed below that found in either single agent economy. Notice that we see the characteristic increase in the P/D ratio as the single agent economy grows more slowly due to lower consumption dividends being offset by even lower discount rates.

5   Concluding Remarks

We extend the transform analysis in Duffie et al. (2000) to compute a general class of nonlinear moments for affine jump diffusions, which has a wide range of applications in economics and finance. We illustrate the power of this method with examples from several areas including option pricing, term structure, credit risk modeling, and GMM estimations. We also illustrate the application of the generalized transform method in three in-depth examples: a model of defaultable bond with recovery risk; an equilibrium model of a Lucas economy with non-i.i.d. trees; and a general class of difference-of-opinion models.

A   Proofs

A.1   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 $s \in \mathbb{R}$ (see their Definition 1.)

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

$$\frac{dF}{dP} = \frac{e^{-\int_0^T r_\tau d\tau} e^{\alpha \cdot X_T}}{E_0[e^{-\int_0^T r_\tau d\tau} e^{\alpha \cdot X_T}]}, \tag{83}$$

> 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 a fact a martingale. (Recall, A, B are solutions to an ODE/BVP given in $(8-9)$.)

Assumption 2 is analogous to 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.

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 $(\hat{f}_{\beta \cdot X_T}^F \check{)} = f_{\beta \cdot X_T}^F$ (see Corollary 8.28 in Folland (1984).) Applying this,

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

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. $\square$

A.2 Proof of Proposition 1

(To be completed)

A.3 Proof of Proposition 2

(To be completed)

B Computations

We wish to compute expectations of the form $E_0[f(X_t)g(\alpha \cdot X_t)]$. Here f will be *pl-linear*, for example $f(x) = e^{a \cdot X_t}(b + c \cdot X_t)$. g is a general function defined on the real line.

In the case of defaultable bonds,

$$
\begin{aligned}
E_0[f(X_t)g(\beta \cdot X_t)] &= \frac{1}{2\pi} E_0[f(X_t) \int_{s \in \mathbb{R}} \hat{g}(s) e^{is\beta \cdot X_t} ds] \\
&= \frac{1}{2\pi} \int_{s=-\infty}^{\infty} E_0[f(X_t) e^{is\beta \cdot X_t}] \hat{g}(s) ds
\end{aligned}
$$

and g is some state-dependent recovery as fraction of face. Here:

$$
\begin{aligned}
f(X_t) &= e^{-\int_0^t (r_s + \lambda_s) ds} \lambda_t \\
&= e^{-\int_0^t (\tilde{\rho}_0 + \tilde{\rho}_1 \cdot X_t) ds} (\lambda_0 + \lambda_1 \cdot X_t)
\end{aligned}
$$

And we can compute:

$$
E_0[e^{-\int_0^t (\tilde{\rho}_0 + \tilde{\rho}_1 \cdot X_s) ds + a + b \cdot X_t}] = e^{A(t;\Theta) + B(t;\Theta) \cdot X_0}
$$

$$
\begin{aligned}
\frac{d}{dt} B &= K_1^\top B + \frac{1}{2} B^\top H_1 B - \tilde{\rho}_1 \quad B(0) = b \\
\frac{d}{dt} A &= K_0^\top B + \frac{1}{2} B^\top H_0 B - \tilde{\rho}_0 \quad A(0) = a
\end{aligned}
$$

Also, considering

$$
h(x) = E_0[e^{-\int_0^t (\tilde{\rho}_0 + \tilde{\rho}_1 \cdot X_s) ds + a + b \cdot X_t + x(c + d \cdot X_t)}]
$$

we can differentiate to obtain

$$
h'(0) = E_0[e^{-\int_0^t (\tilde{\rho}_0 + \tilde{\rho}_1 \cdot X_s) ds + a + b \cdot X_t} (c + d \cdot X_t)].
$$

And so by differentiating the ODE for $h(x)$ and setting $x = 0$ we obtain:

$$E_0[e^{-\int_0^t (\tilde{\rho}_0 + \tilde{\rho}_1 \cdot X_s)ds + a + b \cdot X_t}(c + d \cdot X_t)] = e^{A(t) + B(t) \cdot X_0}(C(t) + D(t) \cdot X_0)$$

$$\frac{d}{dt}B = K_1^\top B + \frac{1}{2}B^\top H_1 B - \tilde{\rho}_1 \quad B(0) = b$$

$$\frac{d}{dt}A = K_0^\top B + \frac{1}{2}B^\top H_0 B - \tilde{\rho}_0 \quad A(0) = a$$

$$\frac{d}{dt}D = K_1^\top D + \frac{1}{2}D^\top H_1 B \quad D(0) = c$$

$$\frac{d}{dt}C = K_0^\top C + \frac{1}{2}D^\top H_0 B \quad C(0) = d$$

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, Dilip Madan, and Frank Zhang. Understanding the role of recovery in default risk models: Empirical comparisons and implied recovery rates. Working paper, 2006.

John Y. Campbell, Jens Hilscher, and Jan Szilagyi. In search of distress risk. *Journal of Finance*, 63(6):2899–2939, 2008.

Mark S. Carey and Michael Gordy. The bank as grim reaper: debt composition and recoveries on defaulted debt. *Proceedings*, (May):336–354, 2007.

Hui Chen. Macroeconomic conditiosn and the puzzles of credit spreads and capital structure. Working Paper, MIT, 2008.

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 Jean Helwege. Is event risk priced? modeling contagion via the updating of beliefs. Working Paper, 2002.

Pierre Collin-Dufresne, Robert Goldstein, and Julien Hugonnier. A general formula for valuing defaultable securities. *Econometrica*, 72(5):1377–1407, 2004.

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.

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.

J. Darrell Duffie, Leandro Saita, and Ke Wang. Multi-period corporate default prediction with stochastic covariates. *Journal of Financial Economics*, 83(3):635–665, 2007.

Bernard Dumas, Alexander Kurshev, and Raman Uppal. Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility. Forthcoming, Journal of Finance, 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.

Robert A. Jarrow and Fan Yu. Counterparty risk and the pricing of defaultable securities. *Journal of Finance*, 56(5):1765–1799, 2001.

Scott Joslin. Pricing and hedging volatility in fixed income markets. Working paper, MIT, 2009.

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.

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.

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.