

Online Appendix for Generalized Transform Analysis of Affine Processes and Applications in Finance.

Hui Chen* Scott Joslin†

September 20, 2011

These notes supplement [Chen and Joslin \(2011\)](#).

1 Heterogeneous agents economy

The generalized transform can also be used to solve models with heterogeneous agents. The heterogeneity could be about preferences or beliefs. Earlier work include [Dumas \(1989\)](#) and [Jiang \(1996\)](#), among others. The stochastic discount factors in the general form of these models will be implicit functions of stochastic state variables. However, even in these general cases, the generalized transform can still be applied.

For illustration, we assume that there is a single risky asset in an endowment economy. Two infinitely-lived agents A and B have time-separable preferences over consumption stream $\{c_t\}$:

$$U_i(c) = E_0 \left[\int_0^\infty u_i(c_t, t) dt \right], \quad i = A, B \quad (\text{O1})$$

We model the dividend process D from the risky asset as part of a general affine jump-diffusion. Specifically, suppose log dividend $d_t = \iota_1 \cdot X_t$, where X_t is a vector that follows the process (7). This model can easily capture features such as predictability in dividend growth, stochastic volatility, or

*MIT Sloan School of Management and NBER, huichen@mit.edu. Tel. 617-324-3896.

†USC Marshall School of Business, sjoslin@usc.edu. Tel. 213-740-7137.

time-varying probability of jumps. For example, to get predictability of dividend growth, we can assume

$$dd_t = g_t dt + \sigma_d dW_t^d \quad (\text{O2a})$$

$$dg_t = \kappa_g (\bar{g} - g_t) dt + \sigma_g dW_t^g \quad (\text{O2b})$$

where g_t is the expected growth rate of dividends, which follows an Ornstein-Uhlenbeck process with long run mean \bar{g} . The state variable is then given by $X_t = [d_t \ g_t]'$, and it is straightforward to verify that X is affine with coefficients implied by the dynamics of d_t and g_t . Finally, as discussed in [Section 3.2](#), other non-AJD processes with tractable conditional characteristic functions can be used for X_t as well.

Assuming markets are complete, we can first solve for the optimal allocations through the social planner's problem

$$\max_{\{C_A, C_B\}} E_0 \left[\int_0^\infty \{u_A(C_{A,t}, t) + \lambda u_B(C_{B,t}, t)\} dt \right], \quad (\text{O3})$$

which has constant relative Pareto weight λ and is subject to the market clearing condition $C_{A,t} + C_{B,t} = C_t$. For concreteness, consider the case where the agents have power utility and differ only in their relative risk aversion, $u_i(c, t) = e^{-\rho t} c^{1-\gamma_i} / (1-\gamma_i)$, with $\gamma_A \neq \gamma_B$. The optimal allocations can be solved through the planner's first order conditions and the market clearing condition. Except for a few special cases,¹ agent A's equilibrium consumption is an implicit function of aggregate endowment

$$C_{A,t} = f \left(D_t^{1-\frac{\gamma_B}{\gamma_A}} \right) D_t. \quad (\text{O4})$$

Then, the unique stochastic discount factor in this economy is given by

$$\xi_t = e^{-\rho t} f \left(D_t^{1-\frac{\gamma_B}{\gamma_A}} \right)^{-\gamma_A} D_t^{-\gamma_A} = e^{-\rho t} g(d_t) e^{-\gamma_A d_t}. \quad (\text{O5})$$

where g is an implicit function that is smooth and bounded. The generalized transform can now be applied when we use the discount factor to price claims. For example, the price of a zero coupon

¹[Jiang \(1996\)](#) shows that the model can be solved in closed form when $\gamma_A = n\gamma_B$ where $n = 2, 3, 4$.

bond that pays one unit of consumption at time T is

$$B(t, T) = E_t \left[\frac{\xi_T}{\xi_t} \right] = \frac{e^{-\rho(T-t)}}{\hat{f}(d_t)e^{-\gamma_A d_t}} E_t \left[e^{-\gamma_A \iota_1 \cdot X_T} g(\iota_1 \cdot X_T) \right], \quad (\text{O6})$$

where [Theorem 1](#) can be applied to evaluate the expectation in [\(O6\)](#). In case where there is an explicit formula for g , the transform may be known in closed form. However, even when g does not have a closed form solution, one will only need to compute \hat{g} numerically once: this single calculation can then be used for the valuation of a variety of securities and need not be re-computed for different horizons.

Besides heterogeneous preferences, the above model framework can also be used to study heterogeneity in beliefs across agents, provided that the beliefs of the agents satisfy the “affine-disagreement framework” ([Hui Chen and Joslin \(2010\)](#)).² We can also extend the model to $N > 2$ agents, which will only change the functional form of g for the SDF in [\(O5\)](#) (the transform remains one-dimensional), or to a model of international finance with multiple goods and multiple countries as in [Pavlova and Rigobon \(2007\)](#).

2 An Affine Model of External Habit

The external habit model of [Campbell and Cochrane \(1999\)](#) is a workhorse in asset pricing that helps generate a high and time-varying equity premium even though consumption growth is i.i.d. and has low volatility. Solving this model as well as estimation and forecasting can be challenging due to the complicated dynamics of the external habit process. In this example, we construct a habit process based on affine state variables that captures the desired features of the habit model.

Our construction is based on the continuous-time version of the external habit model in [Santos and Veronesi \(2010\)](#). In an endowment economy, the representative agent’s utility over consumption stream $\{C_t\}$ is

$$E \left[\int_0^\infty e^{-\rho t} \frac{(C_t - H_t)^{1-\gamma}}{1-\gamma} dt \right], \quad (\text{O7})$$

²See [Section 3](#) for more discussion on affine differences in beliefs. [Bhamra and Uppal \(2010\)](#) provide a recent treatment of heterogeneous preferences and beliefs when the underlying uncertainties are *i.i.d.*

where H_t is the habit level that is positive and strictly below C_t . The log aggregate endowment $c_t = \log(C_t)$ follows the process

$$dc_t = \mu_c dt + \sigma_c dW_t, \quad (\text{O8})$$

with constant expected growth rate μ_c and volatility σ_c , and we specify the process for H_t later.

The stochastic discount factor is obtained from the marginal utility of consumption for the representative agent,

$$m_t = e^{-\rho t} (C_t - H_t)^{-\gamma} = e^{-\rho t - \gamma c_t} \left(\frac{C_t - H_t}{C_t} \right)^{-\gamma}, \quad (\text{O9})$$

where we rewrite the SDF as the product of the standard SDF for CRRA utility ($e^{-\rho t - \gamma c_t}$) multiplied by a function of the surplus-consumption ratio.

[Campbell and Cochrane \(1999\)](#) specify a heteroskedastic AR(1) process for the surplus-consumption ratio. [Santos and Veronesi \(2010\)](#) directly model $G_t \equiv \left(\frac{C_t - H_t}{C_t} \right)^{-\gamma}$ with a non-affine process that is mean-reverting with stochastic volatility. Our approach is different. Instead of modeling the surplus-consumption ratio or G_t directly, we assume that G_t is a function of a ‘‘habit factor’’ y_t , $G_t = g(y_t)$, where y_t is stationary and jointly affine with c_t . An example for the process of y_t is:

$$dy_t = \kappa(\bar{y} - y_t)dt - \sigma_y dW_t, \quad (\text{O10})$$

where y_t and c_t are instantaneously perfectly negatively correlated, which captures the property that the habit level is solely driven by consumption shocks. The SDF then becomes

$$m(t, X_t) = e^{-\rho t - \gamma c_t} g(y_t). \quad (\text{O11})$$

One can apply several criteria in choosing the functional form for g , which in turn implies the dynamics of G_t . First, we need $g(y) > 1$, because H_t should be between 0 and C_t . Second, we also need $g'(y) \geq 0$, because a negative shock to consumption raises the consumption-surplus ratio and hence the marginal utility of consumption. Third, a negative shock to consumption ought to reduce

the habit level, that is, $\text{Cov}_t(dC_t, dH_t) > 0$. Fourth, to generate counter-cyclical risk premium, we would like negative shocks to consumption to raise the conditional Sharpe ratio of the market portfolio (the price of risk for consumption shocks):

$$SR(y_t) = \gamma\sigma_c + \frac{g'(y_t)}{g(y_t)}\sigma_y. \quad (\text{O12})$$

Thus, we need $\frac{d}{dy}SR(y) = \frac{d}{dy}\log(g(y)) \geq 0$.

Finding such a function g is straightforward, and it is clear that it cannot be exponential-affine in y .³ Suppose the desired range of Sharpe ratio is between $\gamma\sigma_c$ and $\gamma\sigma_c + \alpha$ for some $\alpha > 0$. Then, we can assume

$$\frac{g'(y)}{g(y)} = \frac{\alpha}{\sigma_y}F(y), \quad (\text{O13})$$

where F is any monotone non-decreasing function with

$$\lim_{y \rightarrow -\infty} F(y) = 0, \quad \lim_{y \rightarrow +\infty} F(y) = 1.$$

For example, F can be the cumulative distribution function of any real univariate random variable. Thus, we have a lot of flexibility in choosing the desired shape of F , which in turn decides the distribution of the conditional Sharpe ratio.

It follows that

$$g(y) = \exp\left(b + \frac{\alpha}{\sigma_y} \int_{-\infty}^y F(t)dt\right), \quad (\text{O14})$$

which satisfies the criteria we discussed earlier for g provided the constant b is sufficiently large.⁴ The SDF in (O11) fits the moment function in Theorem 1 (with an appropriate choice of factorization as in (16)). This example highlights the power of the generalized transform: rather than specifying a complicated process for g directly, we can utilize the flexibility in choosing $g(y)$ for some tractable

³In the special case where $m(t, X_t)$ follows the exponential-affine form in ??, one can show that if consumption shocks are i.i.d. and X_t is affine, the price of consumption shock has to be constant.

⁴To satisfy the condition $\text{Cov}_t(dC_t, dH_t) > 0$, it suffices to have $b > \gamma \ln\left(1 + \frac{\alpha}{\gamma\sigma_c}\right)$.

process y (such as an affine process).

3 Affine Differences in Beliefs

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

3.1 General Setup

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

$$dX_t = \mu_t^A dt + \sigma_t^A dW_t^A + dZ_t^A, \quad (\text{O15})$$

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

Agent B has heterogeneous beliefs which we shall suppose are equivalent. A broad class⁵ of such equivalent beliefs can be characterized as follows. There exists some vector a such that Agent B believes X follows an affine jump-diffusion satisfying

$$dX_t = \mu_t^B dt + \sigma_t^B dW_t^B + dZ_t^B, \quad (\text{O16})$$

where

⁵More generally, we could consider beliefs of the form $e^{h(x_t) - \int_0^t e^{-h(x_s)} \mathcal{D}^1 e^{h(x_s)} ds}$. Provided the integral term remains tractable, the same analysis applies. Compare also the discussion of essentially affine difference of opinions.

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

This difference in beliefs generates a disagreement about not only the drifts of the state variables, but also the jump frequency and the distribution of jump size.⁶

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

$$\eta_t = E_t \left[\frac{dP^B}{dP^A} \right] = \exp \left(a \cdot X_t - \int_0^t \left(a \cdot \mu_s^A + \frac{1}{2} \|\sigma_s^A a\|^2 + \lambda_s^A (\phi_\nu(a) - 1) \right) ds \right). \quad (\text{O17})$$

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

While we specify the differences in beliefs exogenously, this does not preclude agents' beliefs from arising through Bayesian updating based on different information sets. For example, when the state variables and signals follow a joint Gaussian process, Bayesian updating can reduce to a difference of beliefs in the form of (O17).

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

$$\eta_t = e^{a \cdot X_t}. \quad (\text{O18})$$

We assume that the agents have time separable preferences:

$$U^i(c) = E_0^i \left[\int_0^\infty u^i(c_t, t) dt \right], \quad i = A, B. \quad (\text{O19})$$

⁶To be precise, as a process $Z^A = Z^B$ (i.e. the functions $Z^i : \Omega \times [0, \infty) \rightarrow \mathbb{R}$ are the same). Agents disagree about the probability measures on Ω .

Suppose also that

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

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

$$u_c^A(C_t^A, t) = \zeta^A \xi_t, \quad (\text{O20})$$

$$u_c^B(C_t^B, t) = \zeta^B \eta_t^{-1} \xi_t, \quad (\text{O21})$$

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

Market clearing then implies

$$C_t = (u_c^A)^{-1}(\zeta^A \xi_t) + (u_c^B)^{-1}(\zeta^B \eta_t^{-1} \xi_t), \quad (\text{O22})$$

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

$$\xi_t = e^{-\rho t} \left[\left(\frac{1}{\zeta^A} \right)^{1/\gamma} + \left(\frac{\eta_t}{\zeta^B} \right)^{1/\gamma} \right]^\gamma C_t^{-\gamma}. \quad (\text{O23})$$

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

$$\xi_t = e^{-\rho t} g(a \cdot X_t) e^{-\gamma c \cdot X_t}. \quad (\text{O24})$$

With the stochastic discount factor in this form, we may price any asset with *pl-linear* payoffs, such

as bonds and dividend claims, using [Theorem 1.7](#). Our method also applies when the two agents have different risk aversion (γ_A and γ_B). In that case, we can still express $h(c_t, \eta_t)$ in the separable form as in [\(O24\)](#), and proceed the same way.

In some cases, the mapping of a difference-of-opinion model to the standard setting [\(O17\)](#) is not immediate, and requires a careful choice of the state variable X_t . For example, consider the setting where the agents believe that (de-trended) aggregate log consumption, c_t , follows an Ornstein-Uhlenbeck process:

$$dc_t = \kappa_A(\theta_A - c_t)dt + \sigma dW_t^A, \quad (\text{O25})$$

$$dc_t = \kappa_B(\theta_B - c_t)dt + \sigma dW_t^B. \quad (\text{O26})$$

In this case, the difference in beliefs cannot be expressed as in [\(O17\)](#) directly. However, by considering an augmented state variable we can return to this form. The state variable $\langle c_t, c_t^2 \rangle$ follows the process

$$d \begin{bmatrix} c_t \\ c_t^2 \end{bmatrix} = \begin{bmatrix} \kappa_A(\theta_A - c_t) \\ 2c_t\kappa_A(\theta_A - c_t) + \frac{1}{2}\sigma^2 \end{bmatrix} dt + \begin{bmatrix} \sigma \\ 2c_t\sigma \end{bmatrix} dW_t^A. \quad (\text{O27})$$

Since the corresponding 2×2 conditional covariance matrix, $[\sigma, 2c_t\sigma]^\top [\sigma, 2c_t\sigma]$, is affine in $\langle c_t, c_t^2 \rangle$, it follows that $\langle c_t, c_t^2 \rangle$ is an affine process. Moreover, we return to our standard case since P^B is given by the change of measure as in [\(O17\)](#) with

$$a = \sigma^{-2} \begin{bmatrix} \kappa_B\theta_B - \kappa_A\theta_A \\ \frac{1}{2}(\kappa_A - \kappa_B) \end{bmatrix}. \quad (\text{O28})$$

More generally, we can have the case where each agent believes that the state of the economy is summarized by the N -dimensional Gaussian state variables, X_t , and each agent believes that X_t

⁷The 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.

satisfies the stochastic differential equation $dX_t = (K_0^i + K_1^i X_t)dt + \sqrt{H_0}dW_t^i$. Again by considering an augmented state variable of the form $\hat{X}_t = \langle X_t, \text{vech}(X_t X_t^\top) \rangle$ we can return to our standard setting.⁸ Such techniques are common in the term structure literature with respect to affine and quadratic term structure models. The procedure generalizes to accommodate models with stochastic volatility ($A_M(N)$ in the parlance of [Dai and Singleton \(2000\)](#)). Following [Duffee \(2002\)](#), we refer to this as *essentially affine difference of beliefs*.

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

$$\lambda_t = \sqrt{H_0^{-1}}(\mu_t^B - \mu_t^A) , \quad (\text{O29})$$

$$\eta_t = e^{-\int_0^t \lambda_s dW_s^A - \frac{1}{2} \int_0^t \|\lambda_s\|^2 ds} . \quad (\text{O30})$$

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

3.2 Special Cases

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

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

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

⁸For a square matrix M , vech denotes the lower triangular entries written as a vector. Usually, only part of the elements in the extended state vector is needed to maintain the Markov structure.

Agent B's probability measure,

$$d \begin{bmatrix} c_t \\ \hat{f}_t^B \\ \hat{g}_t \end{bmatrix} = \begin{bmatrix} \hat{f}_t^B - \frac{1}{2}\sigma_c^2 \\ \kappa(\bar{f} - \hat{f}_t^B) \\ -\psi\hat{g}_t \end{bmatrix} dt + \begin{bmatrix} \sigma_c & 0 \\ \frac{\gamma_B}{\sigma_c} & 0 \\ \sigma_{\hat{g},c} & \sigma_{\hat{g},s} \end{bmatrix} dW_t^B. \quad (\text{O31})$$

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

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

where

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

It is easy to check that the local conditional variance of $X_t, \Sigma_t \Sigma_t^\top$, is affine in X_t so this represents an affine process.⁹ Then, it is immediate that η_t takes the form of (O17) with $a = \langle 0, 1, 0, 0, 0 \rangle$.

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

⁹DKU 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.

density. The formula (A58-A61) in DKU is essentially

$$E_0[e^{\alpha \cdot X_t} g(\beta \cdot X_t)] = \frac{1}{2\pi} \int_{b \in \mathbb{R}} \hat{g}(b) \int_{s \in \mathbb{R}} e^{ibs} E_0[e^{(\alpha - is\beta) \cdot X_t}] ds db. \quad (\text{O32})$$

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

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

$$d \begin{bmatrix} c_t \\ V_t \end{bmatrix} = \begin{bmatrix} \bar{g} \\ -\kappa_V V_t \end{bmatrix} dt + \sqrt{\begin{bmatrix} \sigma_d & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \sigma_{cV} V_t & 0 \\ 0 & \sigma_{VV} V_t \end{bmatrix}} dW_t^A. \quad (\text{O33})$$

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

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

$$d \begin{bmatrix} c_t \\ V_t \end{bmatrix} = \begin{bmatrix} \bar{g} \\ -(\kappa_V - b)V_t \end{bmatrix} dt + \sqrt{\begin{bmatrix} \sigma_d & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \sigma_{cV} V_t & 0 \\ 0 & \sigma_{VV} V_t \end{bmatrix}} dW_t^B. \quad (\text{O34})$$

For example, when $b > 0$, Agent B believe that volatility mean reverts more slowly. Using $a = \langle 0, b/\sigma_{VV}^2 \rangle$ we get the *db-density* as in (O17).

Disagreement about momentum. Consider a model with stochastic growth in consumption. Let c_t be the log consumption, g_t be the expected growth rate. Also, let e_t be an exponential

weighted moving average of past growth rates:

$$e_t = \int_{-\infty}^t e^{-b(t-s)} g_s ds. \quad (\text{O35})$$

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

$$d \begin{bmatrix} c_t \\ g_t \\ e_t \end{bmatrix} = \begin{bmatrix} g_t \\ \kappa(\bar{g} - g_t) \\ g_t - be_t \end{bmatrix} dt + \begin{bmatrix} \sigma_c & 0 \\ 0 & \sigma_g \\ 0 & 0 \end{bmatrix} dW_t^A. \quad (\text{O36})$$

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

$$d \begin{bmatrix} c_t \\ g_t \\ e_t \end{bmatrix} = \begin{bmatrix} g_t + ce_t \\ \kappa(\bar{g} - g_t) \\ g_t - be_t \end{bmatrix} dt + \begin{bmatrix} \sigma_c & 0 \\ 0 & \sigma_g \\ 0 & 0 \end{bmatrix} dW_t^B. \quad (\text{O37})$$

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

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

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

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

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

We can consider the state variable $X_t = [c_t, \hat{g}_t^A, \hat{g}_t^B, \hat{g}_t^{AB}, \hat{g}_t^{BA}]$. Suppose that X_t follows a

Gaussian process under both agents beliefs. Agent A's beliefs are such that

$$d \begin{bmatrix} c_t \\ \hat{g}_t^A \\ \hat{g}_t^B \\ \hat{g}^{AB} \\ \hat{g}^{BA} \end{bmatrix} dt = \begin{bmatrix} \hat{g}_t^A \\ \kappa_A(\theta - \hat{g}_t^A) \\ \kappa_B(\theta - \hat{g}_t^B) \\ \kappa_{AB}(\hat{g}_t^B - \hat{g}^{AB}) \\ \kappa_{BA}(\hat{g}_t^A - \hat{g}^{BA}) \end{bmatrix} dt + \Sigma dW_t^A. \quad (\text{O38})$$

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

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

Disagreement about the likelihood of disasters. Suppose that log consumption, c_t , has constant growth with IID innovations with time-varying probability, λ_t , of rare disaster. Let $X_t = [c_t, \lambda_t]$. Under Agent A's beliefs,

$$dX_t = \begin{bmatrix} g_A \\ -\kappa_\lambda \lambda_t \end{bmatrix} dt + \begin{bmatrix} \sigma_c & 0 \\ 0 & \sigma_\lambda \sqrt{\lambda} \end{bmatrix} dW_t^A + dZ_t^A, \quad (\text{O39})$$

where Z_t^A are jumps in c_t which occur with intensity $\lambda_0 + \lambda_t$ and distribution ν . Suppose that Agent B's beliefs are specified by the db-density of form (O17) with $a = \langle b, 0 \rangle$. Then, Agent B's beliefs will be

$$dX_t = \begin{bmatrix} g_A + b\sigma_c^2 \\ -\kappa_\lambda \lambda_t \end{bmatrix} dt + \begin{bmatrix} \sigma_c & 0 \\ 0 & \sigma_\lambda \sqrt{\lambda} \end{bmatrix} dW_t^B + dZ_t^B, \quad (\text{O40})$$

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

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

γ	6	$\bar{\rho}_{\ell,d}$	-30.3%
ρ	1%	$\bar{\sigma}_{\ell}$	5.4%
\bar{g}	1.5%	$\bar{\sigma}_d$	11.1%
\bar{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

Table 1: **Parameters.** 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 be one.

being a disaster.

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 1](#) 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 $(\Sigma_0$ and $\Sigma_1)$ are difficult to directly interpret, we calibrate them by considering their effect on the volatility of labor income, the 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-risk free rate puzzle ([Mehra and Prescott \(1983\)](#)), and we choose our parameterization to generate 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 $\sigma_{\infty}(\sigma_d) = 1.7\text{bp}$, $\sigma_{\infty}(\sigma_{\ell}) = 0.18\text{bp}$, 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 7 moments (along with the simplifying assumption that innovations to V are

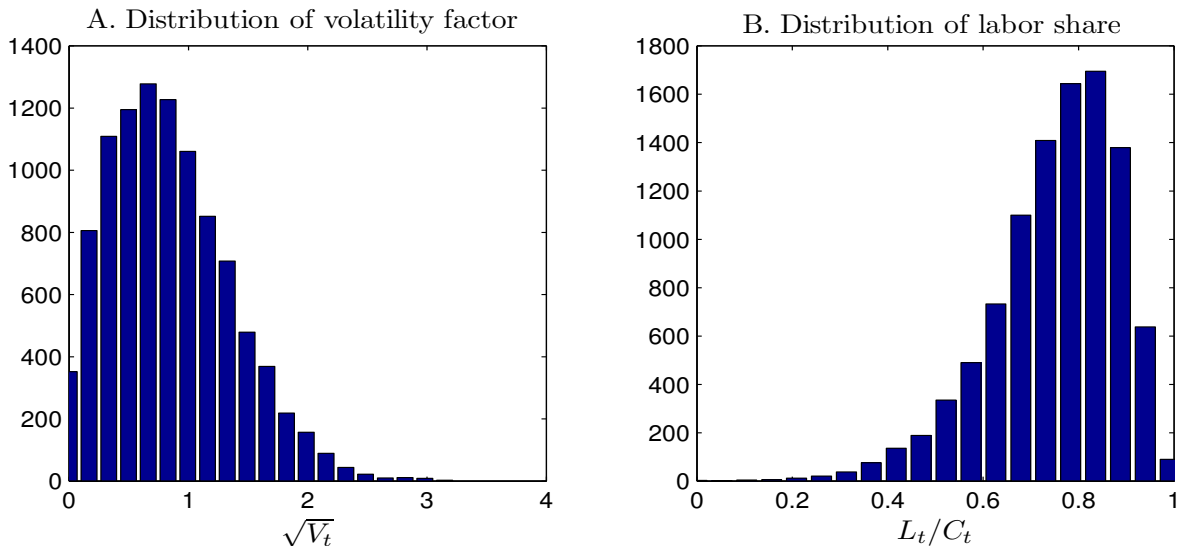


Figure 1: **Simulated distributions.** Panels A and B plot the simulated distributions of the volatility factor $\sqrt{V_t}$ and labor share L_t/C_t .

uncorrelated with innovations to either ℓ or d) fix the free parameters in Σ_0 (3 parameters) and Σ_1 (4 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](#).

Next, we investigate the price of risk for shocks to dividends and labor income, which are transparent and helpful for our understanding of the risk premiums on financial wealth and human capital. From the stochastic discount factor, we can compute the price of dividend risk, which is defined as the risk premium for one unit exposure to dividend shocks:

$$PR_t^d = \frac{\gamma \text{cov}_t(c_t, d_t)}{\sigma_{d,t}} = \gamma \left(\frac{L_t}{C_t} \rho_t \sigma_{\ell,t} + \left(1 - \frac{L_t}{C_t} \right) \sigma_{d,t} \right), \quad (\text{O41})$$

where c_t is log consumption; $\sigma_{\ell,t}$ and $\sigma_{d,t}$ are the conditional volatilities of labor income and dividends, with their dependence on the volatility factor V_t captured by the time subscripts. On the one hand, holding V_t constant (so that $\sigma_{\ell,t}$, $\sigma_{d,t}$, and ρ_t are all constant), the price of dividend risk will rise as labor share falls as long as $\sigma_{d,t} > \rho_t \sigma_{\ell,t}$ (for example, when dividends are more volatile than labor income or when their correlation is negative), which is the composition effect highlighted

in SV. On the other hand, holding labor share fixed, as V_t increases, the volatility of labor income, dividends, and the correlation between the two will increase. All three factors contribute to raise the price of dividend risk. Intuitively, investors become more reluctant to hold financial assets either when there is less labor income to buffer the financial shocks (lower labor share), or when labor income becomes a worse hedge against the financial shocks (higher correlation or higher volatility of labor income).

Finally, if the correlation between labor income and dividends is sufficiently negative and the labor share is sufficiently high, the price of dividend risk can become negative. This is because when the investor is heavily exposed to labor income risk and ρ_t is close to -1 , bad news for dividends will tend to be accompanied by good news to labor income, which causes consumption to rise (opposite to dividends).

Similarly, the price of labor income risk measures the risk premium for one unit exposure to labor income shocks:

$$PR_t^\ell = \frac{\gamma \text{cov}_t(c_t, \ell_t)}{\sigma_{\ell,t}} = \gamma \left(\frac{L_t}{C_t} \sigma_{\ell,t} + \left(1 - \frac{L_t}{C_t} \right) \rho_t \sigma_{d,t} \right). \quad (\text{O42})$$

Provided that $\sigma_{\ell,t} > \rho_t \sigma_{d,t}$ (for example, when the correlation is very small or negative), the price of labor income risk will rise with labor share. Also, holding the labor share constant, the price of labor income risk unambiguously falls as the correlation and volatilities fall.

Figure 2 shows the quantitative effects of the labor share and correlation (volatility) on the price of dividend risk and labor income risk. First, we show how the volatilities of labor income, dividends, and the correlation between the two are tied to the volatility factor V_t . As Panel A shows, both the volatility of labor income and dividends are monotonically increasing in the volatility factor, although dividend volatility varies significantly more than does labor income. This property is an important feature of our model, which is also consistent with the data. In Panel B, the conditional correlation between labor income and dividends is also monotonically increasing in the volatility factor. Since the volatilities and correlation are all monotonic functions of the volatility factor V_t , without loss of generality we use the correlation ρ_t in place of V_t in the plots for the remainder of the paper.

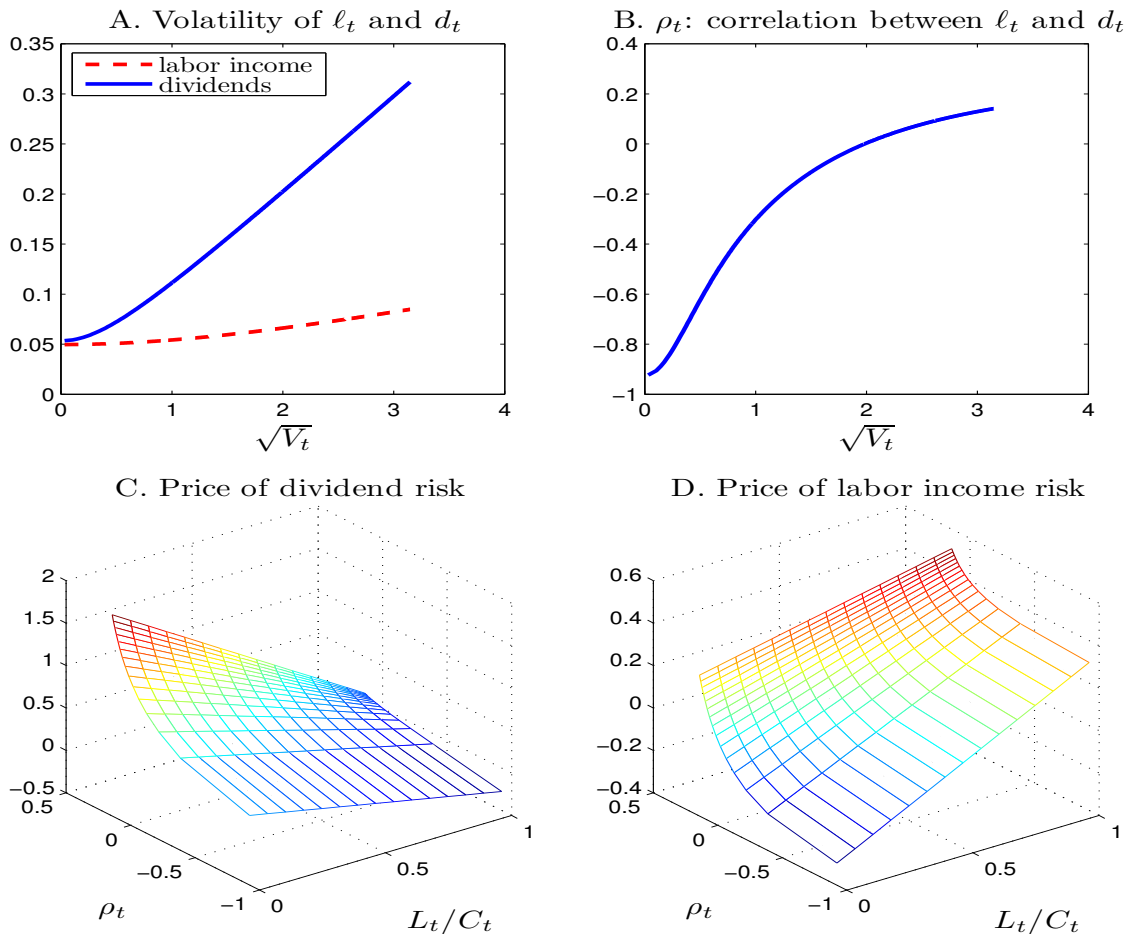


Figure 2: **Price of dividend risk and labor income risk.** Panel A and B plot the conditional volatility of labor income ($\sigma_{\ell,t}$), dividends ($\sigma_{d,t}$), and the conditional correlation between the two (ρ_t), as functions of $\sqrt{V_t}$. Panels C and D plot the conditional price of dividend and labor income risk PR_t^d and PR_t^ℓ as function of labor share L_t/C_t and correlation ρ_t .

As Panels C and D show, holding the correlation constant, the price of dividend risk is decreasing in labor share, while the price of labor income risk is increasing in the labor share. When holding the labor share constant, both prices of risk increase with the correlation between labor income and dividends (the volatility factor). These features are consistent with our analysis above.

Notice also that while the price of dividend risk is always falling as the labor share rises, the decline is less pronounced when the correlation (volatility) is small. For example, when the correlation is -0.1, the price of dividend risk falls from 0.36 to 0.06 as labor share rises from 0.6 to 0.9; when the correlation is -0.8, price of dividend risk falls from 0.01 to -0.17 for the same increase

in labor share. The opposite is true for the price of labor income risk. From (O41) and (O42) we see that the sensitivity of PR_t^d to L_t/C_t is $\sigma_{d,t} - \rho_t\sigma_{\ell,t}$, which is increasing in V_t under our calibration, whereas the sensitivity PR_t^ℓ to L_t/C_t is $\sigma_{\ell,t} - \rho_t\sigma_{d,t}$, which is decreasing in V_t under our calibration. Both results follow from the fact that the volatility of dividend income $\sigma_{d,t}$ rises significantly faster with V_t than does the volatility of labor income $\sigma_{\ell,t}$, as illustrated in Panel A of Figure 2.

Panels C and D also show that the decline in the prices of dividend risk and labor income risk with correlation (and volatility) is more pronounced when labor share is lower. Since $\sigma_{d,t}$ rises significantly faster with V_t than $\sigma_{\ell,t}$ and even $\rho_t\sigma_{\ell,t}$, a lower labor share will make the price of dividend risk more sensitive to changes in the volatilities and correlation between labor income and dividends.

Next, we return to assess whether changing covariance can account for the changing relationship between labor income share and expected excess returns in Figure 3. Specifically, our model predicts that the labor income share has little effect on the equity premium when consumption volatility and the correlation between labor income and dividends are low.

In Table 2, we formally examine the difference of return predictability with labor share in the two samples using regressions of long-horizon excess returns on lagged labor shares. The specific form of the regression is

$$r_{t,t+K}^x = \beta_0(K) + \beta_1(K)\frac{L_t}{C_t} + \epsilon_{t+K}, \quad (\text{O43})$$

where $r_{t,t+K}^x$ is the cumulative excess return on the market over K quarters, and we consider $K = 4, 8, 12, 16$. For each regression, we report the point estimates of $\beta_1(K)$, the NeweyWest corrected t -statistics (with the number of lags equal to $2K$, as in SV), the Hansen-Hodrick corrected t -statistics and the adjusted R^2 .

The regression results are consistent with what we have inferred from Figure 3. Labor share significantly predicts long-horizon returns with a negative sign in the period 1947-1990, which confirms the findings of SV. In the post-1990 period, the regression coefficients on labor share become positive, although they are statistically non-significant at all horizons when using the Hansen-Hodrick standard errors. Based on the moving-window estimates of consumption volatility and correlation between labor income and dividends in Figure 4, volatilities and correlation are both

Table 2: Return Forecasting Regressions. Dependent variable: cumulative excess return on the market over various horizons (quarters). Predictive variable L_t/C_t : share of labor income to consumption. The first set of t -stats is the Newey-West adjusted t -statistics, with the number of lags double the forecasting horizon. The second set of t -stats is the Hansen-Hodrick adjusted t -statistics using the same number of lags as the Newey-West t -stats.

	Panel A: 1947-1990				Panel B: 1991-2010			
	4	8	12	16	4	8	12	16
L_t/C_t	-1.042	-2.440	-4.257	-6.263	3.621	7.233	11.342	15.791
t -stat (NW)	-1.890	-3.664	-4.367	-4.396	1.948	2.185	2.569	3.637
t -stat (HH)	-1.086	-1.401	-1.900	-2.330	1.295	1.406	1.531	1.724
Adj. R^2	0.041	0.125	0.270	0.406	0.101	0.214	0.286	0.363

higher pre-1990, which according to our model implies a stronger composition effect (risk premium falls as labor share rises) during this period. As both the volatilities and correlation become smaller in the second sample, the composition effect should indeed become weaker.

On the other hand, in [Table 2](#) the point estimate of the regression coefficients are economically large and significantly positive at longer horizons with respect to the Newey-West standard errors. What could explain such a possible change in the predictive power of labor share? It is possible that some variable that is driving the risk premium on stocks has become correlated with labor share (or happens to be correlated in the relatively short sample). For example, observe that over this period, labor share and covariance are generally positively correlated in [Figure 4](#). In this case, according to [Figure 5](#), decreasing share may be associated with decreasing risk premium due to declining correlation, despite the reverse univariate relationship where covariance is held fixed. Put differently, our covariance variable, V_t , may represent a variable which is correlated with both the regressor and the residual in [\(O43\)](#), resulting in biased estimates. Other risks factors may also be at play as well. For example, the correlation between labor share and the consumption-wealth measure CAY (see [Lettau and Ludvigson \(2001\)](#)) is -0.15 over the period 1952-2010, but jumps to 0.6 in the period 1990-2010. Thus, the fact that CAY predicts the equity risk premium with a positive sign can help explain the regression results in the post-1990 period as well. These possibilities could be explored further by expanding our framework to include the additional covariates.

References

- Basak, S., 2005, "Asset Prices with Heterogenous Beliefs," *Journal of Banking & Finance*, 29, 2849-2881.
- Bhamra, H., and R. Uppal, 2010, "Asset Prices with Heterogeneity in Preferences and Beliefs," Working Paper.
- Campbell, J. Y., and J. H. Cochrane, 1999, "By Force of Habit: A Consumption Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205-251.
- Chen, H., and S. Joslin, 2011, "Generalized Transform Analysis of Affine Processes and Applications in Finance," Working Paper.
- Cox, J. C., and C. Huang, 1989, "Optimum Consumption and Portfolio Policies When Asset Prices Follow a Diffusion Process," *Journal of Economic Theory*, 49, 33-83.
- Dai, Q., and K. J. Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55, 1943-1978.
- Duffee, G. R., 2002, "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, 57, 405-443.
- Dumas, B., 1989, "Two-Person Dynamic Equilibrium in the Capital Market," *Review of Financial Studies*, pp. 157-188.
- Dumas, B., A. Kurshev, and R. Uppal, 2009, "Equilibrium Portfolio Strategies in the Presence of Sentiment Risk and Excess Volatility," *Journal of Finance*, *Forthcoming*.
- Hui Chen, N.-K. T., and S. Joslin, 2010, "Affine Disagreement and Asset Pricing," *American Economic Review*, 100, 522-526.
- Jiang, W., 1996, "The term structure of interest rates in a pure exchange economy with heterogeneous investors," *Journal of Financial Economics*, 41, 75-110.
- Lettau, M., and S. Ludvigson, 2001, "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying," *Journal of Political Economy*, 109, 1238-1287.

Lettau, M., S. C. Ludvigson, and J. A. Wachter, 2008, “The Declining Equity Premium: What Role Does Macroeconomic Risk Play?,” *The Review of Financial Studies*, Vol. 21, Issue 4, pp. 1653-1687, 2008.

Mehra, R., and E. C. Prescott, 1983, “The Equity Premium: A Puzzle,” *Journal of Monetary Economics*, 15, 145–161.

Pavlova, A., and R. Rigobon, 2007, “Asset Prices and Exchange Rates,” *Review of Financial Studies*, 20, 1139–1181.

Santos, T., and P. Veronesi, 2010, “Habit formation, the cross section of stock returns and the cash-flow risk puzzle,” *Journal of Financial Economics*, 98, 385–413.