# Dynamics of Bond and Stock Returns
# Supplementary Materials: Appendix

**Serhiy Kozak***

December 15, 2021

**Abstract**

A production-based equilibrium model jointly prices bond and stock returns and produces time-varying correlation between stock and real treasury returns that changes in both magnitude and sign. The term premium is time-varying and changes sign. The model incorporates time-varying risk aversion and two physical technologies with different cash-flow risks. Bonds hedge risk-aversion shocks and command negative term premium through this channel. Cash-flow shocks produce co-movement of bond and stock returns and positive term premium. Relative strength of these two mechanisms varies over time. The correlation is a powerful predictor of relative bond-stock and long-short equity returns in the data.

*JEL classification codes:* G11, G12, E21, E23.
*Keywords:* bond-stock correlation, risk premia, general equilibrium.

*R.H. Smith School of Business, University of Maryland. Email: sekozak@umd.edu.

# Supplementary Materials: Appendix

## A Derivations and Proofs

### A.1 Portfolio Problem

Assume complete markets. A representative investor in this economy maximizes his utility over consumption,

$$J_t = \mathbb{E}_t \left( \int_t^T \left[ f\left(C_\tau, J_\tau\right) + \frac{1}{2} A\left(J_\tau\right) \parallel J_{\boldsymbol{X}}\left(\boldsymbol{X_\tau}, \tau\right) \sigma_{\boldsymbol{X}}\left(\boldsymbol{X_\tau}, C_\tau, \tau\right) \parallel^2 \right] d\tau \right), \tag{A.1}$$

subject to total wealth constraint

$$dW_t = \left[ W_t \boldsymbol{\theta_t'} \boldsymbol{\lambda_t} + W_t r_t - C_t \right] dt + W_t \boldsymbol{\theta_t'} \boldsymbol{\sigma_R} d\boldsymbol{Z}, \tag{A.2}$$

where $\boldsymbol{X}\left(\alpha, \boldsymbol{\lambda}\right)$ is a vector of aggregate state variables that are taken by agent as given and evolve according to

$$d\boldsymbol{X} = \mu_{\boldsymbol{X}} dt + \boldsymbol{\sigma_X} d\boldsymbol{Z}, \tag{A.3}$$

and the vector of prices

$$d\boldsymbol{S_t} = \left( \boldsymbol{S_t} \cdot \left[ r_t + \boldsymbol{\lambda_t} \right] - \boldsymbol{D_t} \right) dt + \boldsymbol{S_t} \cdot \boldsymbol{\sigma_R} d\boldsymbol{Z}. \tag{A.4}$$

The flow utility function can be expressed as

$$U\left(C_\tau\right) = f\left(C_\tau, J_\tau\right) - \frac{1}{2} \frac{\alpha_t}{J_t} \| J_W W_t \boldsymbol{\theta_t'} \boldsymbol{\sigma_R} + J_X \boldsymbol{\sigma_X} \|^2. \tag{A.5}$$

The first-order conditions are

$$\begin{aligned} f_C &= J_W \\ 0 &= -\frac{\alpha}{J} \boldsymbol{\sigma_J} \boldsymbol{\sigma_R'} + \boldsymbol{\lambda} + \frac{J_{WW}}{J_W} W \boldsymbol{\sigma_R} \boldsymbol{\sigma_R'} \boldsymbol{\theta} + \boldsymbol{\sigma_R} \boldsymbol{\sigma_X'} \frac{J_{W\boldsymbol{X}}}{J_W}. \end{aligned} \tag{A.6}$$

### A.2 SDF

#### A.2.1 Duffie-Epstein aggregators and the SDF when $\alpha$ is constant

**Ordinally equivalent aggregator** Define the change of variables

$$\begin{aligned} \chi\left(J\right) \equiv \bar{J} &= \frac{1}{1-\alpha} J^{1-\alpha} \\ \chi'\left(J\right) &= J^{-\alpha} \\ \chi''\left(J\right) &= -\alpha J^{-\alpha-1}. \end{aligned} \tag{A.7}$$

Duffie and Epstein (1992b) call two aggregators $(f, A)$ and $\left(\bar{f}, \bar{A}\right)$ ordinally equivalent, if there is a change of variable $\chi$ such that the following two conditions hold:

$$f\left(C, J\right) = \frac{\bar{f}\left(C, \chi\left(J\right)\right)}{\chi'\left(J\right)} \tag{A.8}$$

$$A\left(J\right) = \chi'\left(J\right) \bar{A}\left(\chi\left(J\right)\right) + \frac{\chi''\left(J\right)}{\chi'\left(J\right)}. \tag{A.9}$$

We can now find an ordinally equivalent aggregator $\left(\bar{f},\bar{A}\right)$ produced by the change of variables in Eq. A.7:

$$\begin{aligned}\bar{f}\left(C,\chi\left(J\right)\right) &= \frac{\phi}{\rho}\frac{C^{\rho}-\left(\left(1-\alpha\right)\bar{J}\right)^{\frac{\rho}{1-\alpha}}}{\left(\left(1-\alpha\right)\bar{J}\right)^{\frac{\rho}{1-\alpha}-1}} \qquad (A.10)\\ \frac{-\alpha}{J} &= \frac{\chi''\left(J\right)}{\chi'\left(J\right)} \quad\Longrightarrow\quad \bar{A}=0.\end{aligned}$$

Therefore two aggregators $(f,A)$ and $\left(\bar{f},0\right)$ are ordinally equivalent with a change of variables $\chi\left(J\right)$ defined in Eq. A.7. Furthermore, since $\bar{A}=0$, the aggregator $\left(\bar{f},0\right)$ is a *normalized* aggregator with $\bar{f}\left(C,\bar{J}\right)$ given by Eq. A.10.

We can now use the normalized aggregator to derive the SDF.

**SDF** The SDF is given by Duffie and Epstein (1992b):

$$\frac{d\Lambda}{\Lambda} \equiv \bar{f}_{V}\left(C,\bar{J}\right)dt+\frac{d\bar{f}_{C}\left(C,\bar{J}\right)}{\bar{f}_{C}\left(C,\bar{J}\right)}. \qquad (A.11)$$

The loading on shocks, $\mathcal{L}\left(\frac{d\Lambda}{\Lambda}\right)$ is given by

$$\begin{aligned}\mathcal{L}\left(\frac{d\Lambda}{\Lambda}\right) &= \mathcal{L}\left(dln\bar{f}_{C}\left(C,\bar{J}\right)\right) \qquad (A.12)\\ &= \mathcal{L}\left(dlnf_{C}\left(C,J\right)-\alpha dlnJ\right). \qquad (A.13)\end{aligned}$$

Note this expression is a special case of the equivalent one in section A.2.2 when $\alpha$ is constant.

### A.2.2 SDF when $\alpha$ is time-varying

To find the SDF, I solve the portfolio problem. I assume that agents take the process for $\alpha$ as given exogenously.

Guess that the SDF is of the form

$$\begin{aligned}d\Lambda &= -r\left(\boldsymbol{X}\right)\Lambda dt+\boldsymbol{\sigma_{\Lambda}}d\boldsymbol{Z} \qquad (A.14)\\ \boldsymbol{\sigma_{\Lambda}} &= \Lambda\times\mathcal{L}\left[dlnf_{C}-\alpha_{t}dlnJ\right], \qquad (A.15)\end{aligned}$$

where $\mathcal{L}\left(\cdot\right)$ denotes the vector of loading on the shocks. Substitute the FOC to get

$$\begin{aligned}\boldsymbol{\sigma_{\Lambda}} &= \Lambda\times\mathcal{L}\left[\frac{J_{WW}}{J_{W}}dW+\frac{J_{W\boldsymbol{X}}}{J_{W}}d\boldsymbol{X}-\frac{\alpha}{J}\boldsymbol{\sigma_{J}}d\boldsymbol{Z}\right] \qquad (A.16)\\ \boldsymbol{\sigma_{\Lambda}} &= \Lambda\frac{J_{WW}}{J_{W}}W_{t}\boldsymbol{\theta_{t}'}\boldsymbol{\sigma_{R}}+\Lambda\frac{J_{W\boldsymbol{X}}}{J_{W}}\boldsymbol{\sigma_{X}}-\Lambda\frac{\alpha}{J}\boldsymbol{\sigma_{J}} \qquad (A.17)\\ \boldsymbol{\sigma_{R}}\boldsymbol{\sigma_{\Lambda}'} &= -\boldsymbol{\lambda}\Lambda. \qquad (A.18)\end{aligned}$$

Define $\boldsymbol{Y_{t}}=\boldsymbol{S_{t}}\Lambda_{t}$. Net of dividends, it should be a martingale,

$$\begin{aligned}\boldsymbol{\mu_{Y}} &= -\Lambda_{t}\boldsymbol{D_{t}} \qquad (A.19)\\ -\Lambda_{t}\boldsymbol{S_{t}}r+\Lambda_{t}\left(\boldsymbol{S_{t}}\cdot\left[r+\boldsymbol{\lambda_{t}}\right]-D\right)+\boldsymbol{S_{t}}\boldsymbol{\sigma_{R}}\boldsymbol{\sigma_{\Lambda}'} &= -\Lambda_{t}\boldsymbol{D_{t}} \qquad (A.20)\\ -\boldsymbol{\lambda_{t}}\Lambda_{t} &= \boldsymbol{\sigma_{R}}\boldsymbol{\sigma_{\Lambda}'}. \qquad (A.21)\end{aligned}$$

Hence the guess was indeed correct. Therefore, the SDF is given by

$$\frac{d\Lambda}{\Lambda}=-r\left(\boldsymbol{X}\right)dt+\mathcal{L}\left[dlnf_{C}-\alpha_{t}dlnJ\right]d\boldsymbol{Z}. \qquad (A.22)$$

Note that if $\alpha$ is constant, $\Lambda$ integrates to the usual expression, $\Lambda=\text{const}\times f_{C}J^{-\alpha}$ which is the same as we get when using normalized aggregator (see section A.2.1). If $\alpha$ is not constant, however, the expression above does not integrate easily.

## A.3 Planner's Problem

### A.3.1 Derivation of the PDE

Planner chooses investment and consumption in order to maximize agent's lifetime utility in Eq. 3.

**HJB equation** Define the flow utility $U(C_\tau)$ as by $U(C_\tau) = f(C_\tau, J_\tau) + \frac{1}{2} A(J_\tau) \boldsymbol{\sigma_{J,\tau}} \boldsymbol{\sigma}'_{\boldsymbol{J,\tau}}$ where $\boldsymbol{\sigma_J} = J_1 K_1 \boldsymbol{\sigma_K} + J_\alpha \boldsymbol{\sigma_\alpha}$, $J = J(K_0, K_1, \alpha; t)$ is a continuation value, and $J_n$ denotes a derivative of $J$ with respect to $K_n$. The Hamilton-Jacobi-Bellman (HJB) equation for the planner's problem is given by

$$
\begin{aligned}
0 \quad = \sup_{\{i_0, i_1\}_t} \quad & \Big\{ U(C_t) + J_0 \mathbb{E}(dK_0)/dt + J_1 \mathbb{E}(dK_1)/dt + J_\alpha \mathbb{E}(d\alpha)/dt \\
& + \frac{1}{2} J_{11} \mathbb{E}\left(dK_1^2\right)/dt + \frac{1}{2} J_{\alpha\alpha} \mathbb{E}\left(d\alpha^2\right)/dt + J_{1\alpha} \mathbb{E}(dK_1 d\alpha)/dt \Big\}.
\end{aligned} \tag{A.23}
$$

Due to homogeneity, we can guess that solution is linear in $K$:

$$
J(K_0, K_1, \alpha; t) \equiv J(K_0, K_1, \alpha) = K \times F(x, \alpha). \tag{A.24}
$$

**Value function guess** Find a solution of the form

$$
\begin{aligned}
J(K_0, K_1, \alpha; t) &\equiv J(K_0, K_1, \alpha) = K \times F(x, \alpha) && \text{(A.25)} \\
K J_0 / J &= 1 - \frac{F_x}{F} x && \text{(A.26)} \\
K J_1 / J &= 1 + \frac{F_x}{F}(1 - x) && \text{(A.27)} \\
K^2 J_{11} / J &= \frac{F_{xx}}{F}(1 - x)^2 && \text{(A.28)} \\
J_\alpha / J &= \frac{F_\alpha}{F} && \text{(A.29)} \\
J_{\alpha\alpha} / J &= \frac{F_{\alpha\alpha}}{F} && \text{(A.30)} \\
K J_{1\alpha} / J &= \frac{F_\alpha}{F} + \frac{F_{x\alpha}}{F}(1 - x). && \text{(A.31)}
\end{aligned}
$$

Volatility of the value function $\boldsymbol{\sigma_J}$ is given by

$$
\begin{aligned}
J &= K \times F(x, \alpha) && \text{(A.32)} \\
\boldsymbol{\sigma_J} &= F(x, \alpha) K x \boldsymbol{\sigma_K} + K \times sd(dF) && \text{(A.33)} \\
\boldsymbol{\sigma_J} &= J x \boldsymbol{\sigma_K} + J \frac{1}{F} \times sd(F_x dx + F_\alpha d\alpha) && \text{(A.34)} \\
\boldsymbol{\sigma_J} &= J x \boldsymbol{\sigma_K} + J \frac{1}{F}\left[F_x x(1 - x) \boldsymbol{\sigma_K} + F_\alpha \alpha \boldsymbol{\sigma_\alpha}\right] && \text{(A.35)} \\
\boldsymbol{\sigma_J} &= J\left[\left(1 + \frac{F_x}{F}(1 - x)\right) x \boldsymbol{\sigma_K} + \frac{F_\alpha}{F} \alpha \boldsymbol{\sigma_\alpha}\right] && \text{(A.36)} \\
&\equiv J \boldsymbol{\sigma_F}. && \text{(A.37)}
\end{aligned}
$$

Solution to the problem above is given by a system of one second-order PDE in two state variables, two first-order conditions for optimal investment, and the aggregate budget constraint

$$
\begin{aligned}
& \frac{\delta}{\rho}\left[\left(\frac{c(x, \alpha)}{F(x, \alpha)}\right)^\rho - 1\right] - \frac{1}{2} \alpha \|\boldsymbol{\sigma_F}\|^2 + (1 - x)\left(1 - \frac{F_x}{F} x\right) \phi_0(i_0) + x\left(1 + \frac{F_x}{F}(1 - x)\right) \phi_1(i_1) \\
& + \frac{F_\alpha}{F} \phi(\bar{\alpha} - \alpha) + \frac{1}{2} \frac{F_{xx}}{F}(1 - x)^2 x^2 \varsigma_K^2 + \frac{1}{2} \frac{F_{\alpha\alpha}}{F} \alpha^2 \|\boldsymbol{\sigma_\alpha}\|^2 + \left[\frac{F_\alpha}{F} + \frac{F_{x\alpha}}{F}(1 - x)\right] x \alpha \lambda \varsigma_K^2 = 0
\end{aligned} \tag{A.38}
$$

$$\delta\left(\frac{c}{F}\right)^{\rho-1} = (F - F_x x)\,\phi_0^{'}(i_0) \tag{A.39}$$

$$\delta\left(\frac{c}{F}\right)^{\rho-1} = (F + F_x(1-x))\,\phi_1^{'}(i_1) \tag{A.40}$$

$$c = (A_0 - i_0)(1-x) + (A_1 - i_1)\,x, \tag{A.41}$$

that have to be solved jointly.

### A.3.2 Asset Prices

Marginal $q$ of technologies are

$$q_0 = \frac{1}{\delta}\left[1 - \omega_x(x,\alpha)\,x\right]c \tag{A.42}$$

$$q_1 = \frac{1}{\delta}\left[1 + \omega_x(x,\alpha)(1-x)\right]c \tag{A.43}$$

where $\omega_x(x,\alpha) = \frac{F_x}{F}$, and $\omega_\alpha(x,\alpha) = \frac{F_\alpha}{F}$.

Realized returns on two assets are

$$dR_n = \frac{A_n - i_n}{q_n}dt + \frac{dq_n}{q_n} + \frac{dK_n}{K_n} + \left\langle \frac{dq_n}{q_n}, \frac{dK_n}{K_n} \right\rangle. \tag{A.44}$$

The loadings of returns on shocks are given by $\mathcal{L}(dR_n) = \mathcal{L}\left(\frac{dK_n}{K_n} + \frac{dq_n}{q_n}\right)$ and the excess returns on two technologies, therefore, are

$$\begin{aligned} RX_n = & \Big(\bar{A}(1-x)\,x\boldsymbol{\sigma_K} + \alpha x\boldsymbol{\sigma_K} + (\alpha-1)\,\omega_x(x,\alpha)\,x(1-x)\,\boldsymbol{\sigma_K} \\ & + (\alpha-1)\,\omega_\alpha(x,\alpha)\,\alpha\boldsymbol{\sigma_\alpha}\Big)\Big(l_{n,k} + l_{n,x}x(1-x)\,\boldsymbol{\sigma_K} + l_{n,\alpha}\alpha\boldsymbol{\sigma_\alpha}\Big)^{\top}, \end{aligned} \tag{A.45}$$

where $l_{n,x}(x,\alpha) \equiv \frac{q_{n,x}}{q_n}$ and $l_{n,\alpha}(x,\alpha) \equiv \frac{q_{n,\alpha}}{q_n}$ are the loadings of $\frac{dq}{q}$ on shocks (price effects), $l_{0,k} = (0,0)$ and $l_{1,k} = \boldsymbol{\sigma_K}$ are the loadings of two technologies on capital shocks in their respective capital-accumulation processes ("cash-flow" effects). The first term of the product above shows the loadings of the SDF on shocks (prices of risk) and is derived using Theorem 2.2 in section A.5.2.

The key to characterizing the excess returns and understanding the dynamics of prices is to describe how the functions $\omega_x(x,\alpha)$, $\omega_\alpha(x,\alpha)$, $l_{n,x}(x,\alpha)$, and $l_{n,\alpha}(x,\alpha)$ look like. Although obtaining closed-form solutions for the unknown functions is impossible, Theorem A.2 below and small-noise expansions in section A.4.2 deliver easy-to-analyze expressions that can be used to better understand underlying mechanisms of the model.

**Expected return on risky asset** Adjustment costs are given by

$$\phi(i) = \xi ln\left(1 + \frac{i}{\theta}\right) \tag{A.46}$$

$$\phi'(i) = \frac{\xi}{\theta + i}. \tag{A.47}$$

Marginal $q$ evolves as follows

$$dq = d\left[\frac{\theta + i}{\xi}\right] = \frac{di}{\xi} \tag{A.48}$$

$$\frac{dq}{q} = (\theta+i)^{-1}\mu_i dt + (\theta+i)^{-1}\sigma_i dZ, \tag{A.49}$$

where

$$
\begin{aligned}
\mu_i &= i_x\mu_x + i_g\mu_g + \frac{1}{2}i_{xx}\sigma_x^2 + \frac{1}{2}i_{gg}\sigma_g^2 + i_{xg}\sigma_x\sigma_g && \text{(A.50)}\\
\sigma_i &= i_x\sigma_x + i_g\sigma_g, && \text{(A.51)}
\end{aligned}
$$

and

$$
\begin{aligned}
\mu_x &= x(1-x)\left[\phi_1(i_1) - \phi_0(i_0) - x\sigma_K^2\right] && \text{(A.52)}\\
\sigma_x &= x(1-x)\sigma_K. && \text{(A.53)}
\end{aligned}
$$

Expected return on risky asset,

$$
\frac{1}{dt}\mathbb{E}(dR_n) = \frac{A_n - i_n}{q_n} + (\theta+i)^{-1}\mu_i + \phi_n(i_n) + (\theta+i)^{-1}\sigma_i\sigma_{K_n}. \tag{A.54}
$$

**Interest rate and excess return** Using the previous two results, we can express the interest rate as

$$
r = \frac{1}{dt}\mathbb{E}(dR) - rx, \tag{A.55}
$$

where $rx$ is the excess return and is given by

$$
\begin{aligned}
rx &= -\mathbb{E}\left[\frac{d\Lambda}{\Lambda}\frac{dP}{P}\right] && \text{(A.56)}\\
&= -\frac{1}{dt}\left\langle dlnf_C - \alpha_t dlnJ,\ \theta q\sigma_i + \sigma_{K_n}\right\rangle.
\end{aligned}
$$

### A.3.3 Boundary conditions

$x = 0$ **boundary** FOC for $i_0$ gives

$$
F\phi_0'(i_0) = \delta\left(\frac{c}{F}\right)^{\rho-1}, \tag{A.57}
$$

where $c = A_0 - i_0$. With only one shock, the economy is completely riskless, so risk aversion does not matter, $F(\alpha) \equiv F$. Solve for a constant $F$:

$$
F = \left(\frac{\delta}{\phi_0'(i_0)}(A_0 - i_0)^{\rho-1}\right)^{\frac{1}{\rho}}. \tag{A.58}
$$

Hamilton-Jacobi-Bellman equation:

$$
\begin{aligned}
\frac{\delta}{\rho}\left[\left(\frac{c}{F}\right)^{\rho} - 1\right] + \phi_0(i_0) &= 0 && \text{(A.59)}\\
\frac{\phi_0'(i_0)}{\rho}(A_0 - i_0) - \frac{\delta}{\rho} + \phi_0(i_0) &= 0. && \text{(A.60)}
\end{aligned}
$$

This gives $i_0(0,\alpha) = const$. Set $i_1(0,\alpha) = 0$.

$x = 1$ **boundary** This is the usual problem with one high-risk technology. Solve a corresponding ODE for $i_1(1,\alpha)$.

## A.4 Stylized model

Unfortunately, the planner's problem does not have an analytic solution in general. To better understand the mechanisms at work, specializing the general model above to the case when EIS is equal to 1 and installation function $\phi_n(\cdot)$ is the same for two technologies and takes a log form, is therefore useful. After the discussion of the stylized model, I will solve numerically the *full* model, free of the next two assumptions.

**Assumption A.1.** *Elasticity of intertemporal substitution (EIS) is equal to unity,* $\psi = 1$.

**Assumption A.2.** *Installation function $\phi(i_n)$ is the same for two technologies and takes the log form,*

$$\phi_n(i_n) = \xi \times ln\left(1 + \frac{i_n}{\xi}\right). \tag{A.61}$$

The functional form in Assumption A.2 is concave, ensuring high levels of investment or disinvestment is costly. It has a slope equal to one at $i_n = 0$, i.e., no adjustment costs on the margin at zero investment. I also set depreciation equal to zero. A corresponding adjustment cost function is $\varphi(i) = i - \xi ln\left(1 + \frac{i_n}{\xi}\right)$, which is a convex function.

Supplementary Materials: Appendix, section A.5 specializes the solution in Theorem 2.1 to the case when Assumption A.1 and Assumption A.2 hold.

### A.4.1 Aggregates

Assumption A.1 and Assumption A.2 deliver the following result.

**Theorem A.1.** *Aggregates in the stylized economy are given by*

$$q_t = \frac{A_t + \xi}{\xi + \delta} \tag{A.62}$$

$$i_t = \frac{A_t - \delta}{\xi + \delta} \tag{A.63}$$

$$c_t = \delta \frac{A_t + \xi}{\delta + \xi}, \tag{A.64}$$

*where $q_t = (1 - x_t) q_{0,t} + x_t q_{1,t}$ is the aggregate Tobin's $q$, $i_t = (1 - x_t) i_{0,t} + x_t i_{1,t}$ is the aggregate investment per unit of capital, $c_t$ is the aggregate consumption per unit of capital, and $A_t = (1 - x_t) A_0 + x_t A_1$ is the aggregate productivity.*

*Proof.* Using the fact that $q_n = \frac{1}{\phi'(i_n)} = \frac{\xi}{\xi + i_n}$ and expressions for investment in Eq. A.72, we get $q = (1 - x) q_0 + x q_1 = \frac{1}{\delta} c$. Investment is $i = (1 - x) i_0 + x i_1 = \frac{\xi}{\delta} c - \xi = \xi q - \xi$. Using the resource constraint, we know $i = A - c = A - \delta q$. Equalizing both expressions gives $q = \frac{A+\xi}{\xi+\delta}$. □

**Corollary A.1.** *When elasticity of intertemporal substitution is equal to one, the aggregate consumption-to-wealth ratio is constant, $\frac{C}{W} \equiv \frac{c}{q} = \delta$.*

The aggregates therefore vary in time *only* because the aggregate productivity, $A_t$, is time-varying, $A_t = (1 - x_t) A_0 + x_t A_1$. In fact, if technologies were identical and thus $A_0 = A_1$, all aggregates relative to capital would be constant, and the only reason the levels of aggregate variables vary is because the level of capital varies. This implication and Corollary A.1, however, may be viewed as benefits, because they allow us to consider the pricing implications for two technologies *independently* of those of aggregate economy and to grasp additional economic insights about the underlying mechanisms of the model. I will later solve the model numerically in section 3, without relying on Assumption A.1 and Assumption A.2, and analyze the solution to better understand how both aggregate and relative pricing mechanisms interact.

The Supplementary Materials: Appendix, section A.5.2 shows derivations of expressions for market risk premium and risk-free rate in the economy. Finally, the following result is useful for future analysis.

**Definition A.1.** A *point of equal investment* (PEQ) of the economy in section A.4.1 is an equilibrium in which the level of investment in risky and riskless technologies are equal for a given value of risk aversion $\alpha$, $i_0(x^*(\alpha), \alpha) = i_1(x^*(\alpha), \alpha)$, where $x^* = x^*(\alpha)$ denotes a share of risky capital at a PEQ. Two technologies grow at the same rate at a PEQ.

**Theorem A.2.** *Derivative of the value function $\omega_x(x, \alpha) = \frac{F_x}{F}$ is zero at a PEQ, above zero as $x$ approaches a PEQ from below, $x \nearrow x^*(\alpha)$, and below zero for $x \searrow x^*(\alpha)$.*

*Proof.* When investments are equal, $1 - x^* \frac{F_x^*}{F^*} = 1 + (1 - x^*) \frac{F_x^*}{F^*}$, which implies $F_x(x^*(\alpha), \alpha) = 0$ and thus $\omega_x(x, \alpha) = 0$. Next, $\frac{\partial}{\partial x}\left[\frac{F_x}{F}\right]\Big|_{x=x*} = \frac{F_{xx}^*}{F^*}$. Because no reallocation of capital optimally takes place, the value function must be maximized with respect to $x$, implying $F_{xx}(x^*(\alpha), \alpha) < 0$. □

Although obtaining closed-form solutions for asset prices is impossible in general, the small-noise expansions in the following section deliver easy-to-analyze analytic approximations that can be used to better understand underlying mechanisms of the model.

### A.4.2 Small-noise expansions

Define a perturbation parameter $\epsilon \in [0,1]$ such that when $\epsilon = 0$, the economy is along its deterministic trajectory, and when $\epsilon = 1$, the economy corresponds to the economy of interest. I perturb the process for evolution of risky capital, productivity of risky technology, and risk aversion as follows:

$$dK_1 = \phi_1(i_1) K_1 dt + K_1 \sqrt{\epsilon} \boldsymbol{\sigma_K} d\boldsymbol{Z} \tag{A.65}$$
$$d\alpha = \phi(\bar{\alpha} - \alpha) dt + \sqrt{\epsilon} \boldsymbol{\sigma_\alpha} d\boldsymbol{Z} \tag{A.66}$$
$$A_1(\epsilon) = (1-\epsilon) A_0 + \epsilon A_1. \tag{A.67}$$

Assumption A.1, Assumption A.2, and Corollary A.2 allow us to characterize the solution of the stylized model in terms of a single PDE in Eq. A.72. Moreover, at $\epsilon = 0$, all derivatives of the value function $F$ are zero (in a deterministic steady state, risk aversion has no impact; $x$ is indeterminate because two technologies are riskless and have the same productivity). These two facts simplify computations of expansions around the deterministic path substantially and make computing analytical small-noise expansions feasible.

The perturbation of the productivity of riskless technology in Eq. A.67 is needed because in the steady state in which productivities of two technology are not equal and no uncertainty is present, a technology with lower productivity will be completely dominated by the other technology and thus $x^*$ will be either 0 or 1. Such a steady state might be a very bad point of expansion. Instead of expanding around it, I will seek for expansions around some deterministic path on which the productivities of two technologies are equal.

I therefore proceed with perturbations in three different directions as defined by the system of equations (A.65) – (A.67). An advantage of perturbing around a non-stochastic path is that all expansions of interest can be computed analytically. I first parametrize the value function by a perturbation parameter $\epsilon$, $F(x,\alpha;\epsilon)$. In my notation, $F(x,\alpha;0)$ corresponds to a deterministic path with two equal productivities $A_0 = A_1(\epsilon)|_{\epsilon=0}$, whereas $F(x,\alpha;1)$ corresponds to the value function in the economy of interest. I look for a first-order expansion[1] of $F(x,\alpha;\epsilon)$ as a power series in $\epsilon$:

$$F(x,\alpha;\epsilon) = F(x,\alpha;0) + F_\epsilon(x,\alpha;0)\,\epsilon + o\left(\epsilon^2\right), \tag{A.68}$$

where $F(x,\alpha;0)$ gives the value of $F$ at $\epsilon = 0$ and $F_\epsilon(x,\alpha;\epsilon)$ gives the derivative at $\epsilon = 0$. This derivative, as well as higher-order derivatives, can be computed analytically, which delivers additional insights and characterizations of the behavior around the steady state. Moreover, Judd (1998) emphasizes that whereas the low-order expansions describe the behavior only locally, as we increase the order of expansions, a solution becomes global within the radius of convergence.

Small-noise expansions employed in my paper are closely related to expansions in control-theory literature, namely, Fleming (1971), Fleming and Yang (1994), and James and Campi (1996). Anderson et al. (2012) and Kogan and Uppal (2001) have analyzed a similar type of expansions. These expansions differ from the one typically used in economic literature where expansion takes place in the shock standard deviation and around some deterministic steady state (imposing steady-state values of variables). I expand with respect to the shock variance and around common productivity $A_0$ without imposing steady-state values. The expansion therefore is around some deterministic *trajectory*, rather than a steady state. Moreover, because I am expanding with respect to the shock variance, my first-order expansions correspond to a second-order expansions used in economic literature and my second-order expansions correspond to the fourth-order expansions in the literature (see Anderson et al. (2012)).

I now proceed with summarizing some analytical results of first-order[2] small-noise expansions. All the major mechanisms of the model are operable in these expansions. Note analytical expansions of *any* order can be derived with this method. Higher-order expansions, however, prove to be difficult to analyze and understand, while providing no more of economic intuition.

[1]When $\epsilon \neq 0$, Eq. A.72 is a second-order partial differential equation, but when $\epsilon = 0$, it reduces to a first-order differential equation. This reduction in order induces a so-called "singular perturbation" to the problem, which is often associated with substantial complications. However, as argued by Judd (1998) and formally shown by Fleming (1971), the remarkable feature of stochastic control problems is that perturbation $\epsilon$ can be analyzed as a regular perturbation when it enters as a square root above.

[2]It is often argued that at least second-order perturbations are needed to generate non-zero risk premium, and at least third-order to generate time-variation in risk premium. Because I perturb variance, first order perturbations of a system in Eq. A.66 does generate non-zero risk premia, which also time-vary, due to an exogenous specification of risk aversion.

**Theorem A.3.** *The value function $F(x,\alpha;\epsilon)$ can be expressed as*

$$F(x,\alpha;\epsilon) \quad = \quad f_0 + \left[\zeta_A x \bar{A} - \frac{1}{2}(\zeta_0 + \zeta_\alpha \alpha) x^2 \varsigma_K^2\right] f_0 \epsilon + o\left(\epsilon^2\right), \tag{A.69}$$

*where $f_0 \equiv F(x,\alpha;0)$ is the value function evaluated at $\epsilon = 0$, $\bar{A}$ is defined as $\bar{A} = \frac{A_1 - A_0}{A_0 + \xi}$; $\zeta_A = \frac{\delta+\xi}{\delta}$, $\zeta_0 = \frac{\phi\bar{\alpha}}{\delta(\delta+\phi)}$, and $\zeta_\alpha = \frac{1}{\delta+\phi}$ are constants. Small-noise expansion of the value function around a deterministic trajectory is therefore linear in $\alpha$ and quadratic in $x$.*

*Proof.* Refer to section A.6 for further details. □

To understand the pricing implications for two technologies in the model, I analyze small-noise expansions for Tobin's $q$'s of these technologies (which can be used to infer investment decision rules in a straightforward fashion).

**Theorem A.4.** *Tobin $q$'s of riskless and risky technologies are given by*

$$q_0(x,\alpha;\epsilon) \quad = \quad \zeta_0^q + \zeta_0^q\left[-\frac{\xi}{\delta} x\bar{A} + (\zeta_0 + \zeta_\alpha\alpha) x^2 \varsigma_K^2\right]\epsilon + o\left(\epsilon^2\right) \tag{A.70}$$

$$q_1(x,\alpha;\epsilon) \quad = \quad \zeta_0^q + \zeta_0^q\left[\left(1 + (1-x)\frac{\xi}{\delta}\right)\bar{A} - (\zeta_0 + \zeta_\alpha\alpha) x (1-x) \varsigma_K^2\right]\epsilon + o\left(\epsilon^2\right), \tag{A.71}$$

*where $\zeta_0$, $\zeta_\alpha$, $\zeta_0^q = \frac{A_0+\xi}{\delta+\xi}$ are constants. Expansions of Tobin $q$'s around a deterministic trajectory are therefore linear in $\alpha$ and quadratic in $x$.*

*Proof.* Refer to Eq. A.6 for further details. □

An important feature of small-noise expansions employed in Theorem A.3 and Theorem A.4 is that the value function and prices directly depend on two state variables. We can therefore analyze how shocks to state variables affect each quantity.

Note prices depend on risk aversion $\alpha$. This dependence is an important feature of the model that is *not* present in a single technology model with time-varying risk aversion. In such a model, prices of risk depend on $\alpha$, but the quantity of risk is constant. In the two-technology model, *both* price and quantity of risk are time-varying for each technology, producing potentially significant variation in price and risk premia of underlying technologies and generating interesting joint dynamics.

I use the insights provided by small-noise expansions to establish Propositions A.1 − A.4 below.

**Proposition A.1.** *Around a non-stochastic trajectory of the model in section A.4, shocks to risk aversion that are orthogonal to shocks to the capital-accumulation process move returns on risky and riskless technologies in opposite directions.*

*Proof.* Direct shocks to risk aversion that are uncorrelated with the shock to capital accumulation of risky technology, have only price impacts on the two technologies (via changes in $q$'s). Expressions for prices in Theorem A.4 imply that the price of the riskless technology increases in $\alpha$, while the price of the risky technology falls in $\alpha$. Therefore, a positive shock to $\alpha$ (unexpected increase in risk aversion) increases the contemporaneous return of the riskless technology and lowers the contemporaneous return of the risky technology.

□

**Proposition A.2.** *Around a non-stochastic trajectory of the model in section A.4 and for values of $x$ in $\left[\frac{1}{2}\frac{\xi}{\xi+\delta}x^*,\, 1\right]$, shocks to the risky capital-accumulation process that are orthogonal to shocks to risk aversion move returns on risky and riskless technologies in the same direction. $x^* = x^*(\alpha)$ denotes the value of $x$ at the PEQ.*

*Proof (sketch).* Intuitively, returns on risky technology respond positively to a capital shock. This response is due to the direct loading on the capital shock, which dominates the pricing effect for a broad range of parameter values. In Eq. A.6 I show the returns on the riskless technology respond positively to a capital shock when the share of risky capital $x$ is in the range $\left[\frac{1}{2}\frac{\xi}{\xi+\delta}x^*,\, 1\right]$, where $x^* = x^*(\alpha)$ is the value of $x$ at the PEQ. The rebalancing mechanism drives this positive response. Both returns therefore move in the same direction. Refer to Eq. A.6 for a more formal argument. □

**Proposition A.3.** *Around a non-stochastic trajectory of the model in section A.4, one can find a calibration of the model that produces a positive correlation between returns on riskless and risky technologies for low levels of risk aversion and negative correlation when risk aversion is high.*

*Proof (sketch).* Proposition A.1 and Proposition A.2 establish that two shocks produce different signs of correlation between bond and stock returns. Additionally, in response to a capital shock, discount rates on the risky asset move in a way that dampens the cash-flow effect on the asset. This dampening becomes stronger as risk aversion rises, leading to a relatively weaker comovement of bond and stock returns. At the same time, an increase in risk aversion leads to a stronger "decoupling" of bond and stock returns. As a result, the flight-to-quality mechanism starts dominating the rebalancing mechanism at higher levels of risk aversion. We can therefore select the variance of risk-aversion shocks and parameter $\lambda$ such that the capital effects are stronger for low levels of $\alpha$ and get dominated for high levels of $\alpha$ and hence the proposition holds. Refer to Eq. A.6 for a more formal argument. ☐

**Proposition A.4.** *When two shocks are uncorrelated, the risk premium on risky technology is monotonically increasing in risk aversion for* $x \in \left[0, \min\left(x^* + \frac{\xi}{\xi+\delta}\frac{1}{\bar{A}}x^*, \frac{1}{2}\right)\right]$, *where* $x^* = x^*(\alpha)$ *denotes the value of* $x$ *at the PEQ.*

*Proof.* Refer to Eq. A.6 for more details. ☐

The main purpose of small-noise expansions and the propositions in this section was to analytically characterize the main mechanisms at work and lay the ground for empirical hypotheses I develop in section 4. Although small-noise expansions are valid only sufficiently close to a deterministic trajectory, I verify numerically that higher-order expansions do not overturn the main qualitative results obtained in the section and that the approximation is sufficiently good on the parts of a state space that are visited in equilibrium. Furthermore, I relax Assumption A.1 and Assumption A.2 in section 3 and verify the qualitative results of this section are robust to such modification.

## A.5 Stylized Model: Details

**Corollary A.2.** *When Assumption A.1 and Assumption A.2 hold, the solution to equation in Theorem 2.1 specializes to*

$$\delta\Big[ln(c) - ln(F)\Big] - \frac{1}{2}\alpha\left\|\boldsymbol{\sigma_F}\right\|^2 + (1-x)\left(1-\frac{F_x}{F}x\right)\phi_0(i_0) + x\left(1+\frac{F_x}{F}(1-x)\right)\phi_1(i_1) \tag{A.72}$$

$$+\frac{F_\alpha}{F}\phi(\bar{\alpha}-\alpha) + \frac{1}{2}\frac{F_{\alpha\alpha}}{F}\alpha\left\|\boldsymbol{\sigma_\alpha}\right\|^2 + \frac{1}{2}\frac{F_{xx}}{F}(1-x)^2x^2\varsigma_K^2 + \left[\frac{F_\alpha}{F} + \frac{F_{x\alpha}}{F}(1-x)\right]x\alpha\lambda\varsigma_K^2 = 0,$$

*with investments determined by the first-order conditions,*

$$i_0 + \xi = \frac{\xi}{\delta}\left(1 - x\frac{F_x}{F}\right)c \tag{A.73}$$

$$i_1 + \xi = \frac{\xi}{\delta}\left(1 + (1-x)\frac{F_x}{F}\right)c, \tag{A.74}$$

*and the aggregate resource constraint*

$$c = (A_0 - i_0)(1-x) + (A_1 - i_1)x. \tag{A.75}$$

### A.5.1 Expected returns on the market portfolio

Expected returns on the market portfolio can be readily analyzed in the stylized model,

$$\mathbb{E}(dR_M) = \frac{A-i}{q}dt + \mathbb{E}\frac{dq}{q} + \mathbb{E}\frac{dK}{K} + \left\langle\frac{dq}{q}, \frac{dK}{K}\right\rangle \tag{A.76}$$

$$= \underbrace{\delta + \phi(i_1)}_{\text{1-sector model's } \mathbb{E}R} + \underbrace{\bar{A}\left[\mu_x + (1-x)x^2\varsigma_K^2\right]}_{\mathbb{E}R \text{ due to time-varying productivity}}, \tag{A.77}$$

where $\bar{A} \equiv \frac{A_1 - A_0}{A+\xi}$. The first two terms correspond to a solution of one sector model, as was shown in Example A.1. It comprises of two components: time discounting $\delta$ and expected growth of capital, $\phi(x)$. The last term is new and reflects the presence of time-varying productivity. If two technologies were identical, the term would vanish. Otherwise, it adjusts for growth in the share of high-risk technology $x$, as the economy endogenously reallocates towards its optimal mix of low-risk and high-risk capital via investment.

### A.5.2 Risk prices

Next, I derive the expression for an SDF of the model. Supplementary Materials: Appendix, section A.2.2 shows that risk prices are given by

$$\begin{aligned} -\mathcal{L}\left(\frac{d\Lambda}{\Lambda}\right) &= -\mathcal{L}\left[dlnf_C - \alpha_t dlnJ\right] = \alpha dlnK + dln\tilde{A} + (\alpha-1)\,dlnF \\ &= \underbrace{\alpha x \boldsymbol{\sigma_K}}_{①} + \underbrace{\bar{A}\,(1-x)\,x\boldsymbol{\sigma_K}}_{②} + \underbrace{(\alpha-1)\,\omega_x(x,\alpha)\,x\,(1-x)\,\boldsymbol{\sigma_K}}_{③} + \underbrace{(\alpha-1)\,\omega_\alpha(x,\alpha)\,\alpha\boldsymbol{\sigma_\alpha}}_{④}, \end{aligned} \tag{A.78}$$

where $\tilde{A} = A + \xi$, $\omega_x(x,\alpha) = \frac{F_x}{F}$, and $\omega_\alpha(x,\alpha) = \frac{F_\alpha}{F}$.

The first term is the usual risk price which is also present in a model with one sector. It can be easily seen by setting $x = 1$, $F_x = 0$. In this case the last three terms drop out (when risk aversion is constant) and we end up with an expression for the price of capital risk in one sector economy with constant risk aversion, $rp = \alpha\sigma_K$. See Supplementary Materials: Appendix, section A.5.3 for more details. The second term is a new risk price due to changing productivity. Productivity falls on capital shocks and thus the price of risk is positive. This term is present only due to time-varying productivity; in case when $A_0 = A_1$, it disappears.

The third piece reflects the presence of two distinct technologies. It is present even when productivities of two technologies are equal. It depends on an unknown function $\omega_x(x,\alpha)$. One can show that $\omega_x(x,\alpha) = 0$ at the point where investments in two technologies are equal (PEQ; see the definition below), $\omega_x(x,\alpha) > 0$ when $x$ is below this point, and $\omega_x(x,\alpha) < 0$ when $x$ is above the PEQ. It is therefore positive when we have "too little" capital, and negative otherwise. Risk prices therefore tend to fall as $x$ falls due to direct effect of $x$, but tend to rise due to an increase in $\omega_x(x,\alpha)$. It's not clear which effect dominates.

When risk aversion is time-varying, the last component reflects the appropriate price of risk due to this variation.

**Example A.1.** Aggregates in one-sector economy.

Consider an economy that features only a high-risk sector and constant risk aversion $\alpha$. Eq. A.72 can be easily specialized to this case by setting $x = 1$, $F_x = 0$. For a more formal argument, see Supplementary Materials: Appendix, section A.5.3. In such an economy, all aggregates relative to capital are constant. In particular, $c = \delta\frac{A+\xi}{\delta+\xi}$, $q = \frac{A+\xi}{\xi+\delta}$, and $i = \xi\frac{A-\delta}{\delta+\xi}$. Normalized aggregates in a one-sector and a two-sector economies therefore differ only due to a persistent time-variation in aggregate productivity in a two-sector economy. Supplementary Materials: Appendix, section A.5.3 also shows the expressions for expected market returns, excess returns, and interest rate in a one-sector economy. In particular, I find expected market return is given by $\frac{1}{dt}\mathbb{E}[dR] = \delta + \phi(i)$, excess returns are $rx = \alpha \times \sigma_K^2$ (when risk-aversion is constant), and the risk-free rate is $r = \delta + \phi(i) - \alpha\sigma_K^2$ (when risk aversion is constant).

**Example A.2.** Cox-Ingersoll-Ross (CIR) economy with constant risk aversion.

Consider a model with no adjustment cost, or, equivalently, $\xi = \infty$ in the setup above. When adjusting capital is costless, the economy always supports a constant level of $x \equiv x^* = min\left(\frac{A_1-A_0}{\alpha\sigma_K^2}, 1\right)$. Value function and all aggregates are therefore constant, $c = \delta$, $i = A - \delta$. Marginal $q$ in such an economy is constant and equal to 1 (on aggregate and for each individual technology). All the risk is thus cash-flow risk, $\frac{1}{dt}\mathbb{E}[dR] = (A-i) + i = A$, $\frac{1}{dt}\mathbb{E}[dR_0] = A_0$, $\frac{1}{dt}\mathbb{E}[dR_1] = A_1$, $rx_M = rx_1 = \alpha x^*\sigma_K^2$, $rx_0 = 0$, and $r = A - \alpha x^*\sigma_K^2$. See Supplementary Materials: Appendix, section A.5.4 for derivations.

### A.5.3 One sector model, $EIS = 1$, log adjustment costs, constant $\alpha$

HJB equation:

$$\begin{aligned} \delta\Big[ln(c) - ln(F)\Big] - \frac{1}{2}\alpha\left(\sigma_K + \frac{F_\alpha}{F}\sigma_\alpha\right)\left(\sigma_K + \frac{F_\alpha}{F}\sigma_\alpha\right)' \\ +\phi_1(i) + \frac{F_\alpha}{F}\phi(\bar{\alpha}-\alpha) + \frac{1}{2}\frac{F_{\alpha\alpha}}{F}\sigma_\alpha\sigma_\alpha' + \frac{F_\alpha}{F}\sigma_K\sigma_\alpha' &= 0. \end{aligned} \tag{A.79}$$

Resource constraint:

$$c = A - i. \tag{A.80}$$

Assume log installation function

$$\phi_1(i) = \zeta + \xi \times ln\left(1 + \frac{i}{\xi}\right). \tag{A.81}$$

FOC:

$$\delta \frac{1}{c} = \xi \frac{1}{\xi + i} \tag{A.82}$$

$$c = \frac{\delta}{\xi}(\xi + A - c) \tag{A.83}$$

$$c = \frac{A + \xi}{1 + \frac{\xi}{\delta}} \tag{A.84}$$

$$i = \frac{A + \xi}{1 + \frac{\delta}{\xi}} - \xi \tag{A.85}$$

$$q = \frac{1}{\phi'(i)} = \frac{\xi + i}{\xi} = \frac{A + \xi}{\xi + \delta}. \tag{A.86}$$

So all aggregates $c$, $i$, $q$ (per unit of capital) are constant.

Assume $\sigma_\alpha$ is constant.

Risk premium on this technology can be derived by setting all $F_x = 0$, $x = 1$ in my general two tree model,$rx = \alpha\sigma_K^2$. Expected returns are constant,

$$\frac{1}{dt}\mathbb{E}[dR] = \frac{A - i}{q} + \mathbb{E}\frac{dK}{K} = \delta + \phi(i). \tag{A.87}$$

Risk-free rate is given by

$$r = \frac{1}{dt}\mathbb{E}[dR] - rx = \delta + \phi(i) - \alpha\sigma_K^2. \tag{A.88}$$

So all the variation in risk premium is driven entirely by variation in the risk-free rate.

### A.5.4 CIR Model

Assume constant risk aversion.

$$q = \phi'(i) = 1 \tag{A.89}$$

$$F - F_x x = F + F_x(1 - x) \tag{A.90}$$

$$F_x = 0 \tag{A.91}$$

So value function if flat wrt $x$, since the capital can be freely reallocated. It means that $x$ is not a state variable any more, rather it is a choice: there exists an optimal value $x^*$ which maximizes utility.

HJB equation becomes

$$\frac{\delta}{\rho}\left[\left(\frac{c}{F}\right)^\rho - 1\right] - \frac{1}{2}\alpha x^2\sigma_K^2 + i = 0. \tag{A.92}$$

FOC:

$$\delta\left(\frac{c}{F}\right)^{\rho-1} = F \tag{A.93}$$

$$c = (\delta F^\rho)^{\frac{1}{\rho-1}} \tag{A.94}$$

$$F = \left[(A - i)^{\rho-1}\frac{1}{\delta}\right]^{\frac{1}{\rho}}. \tag{A.95}$$

From HJB,

$$i = \frac{1}{2}\alpha x^2\sigma_K^2 - \frac{\delta}{\rho}\left[\left(\frac{F}{\delta}\right)^{\frac{\rho}{\rho-1}} - 1\right] \tag{A.96}$$

$$c = A - i \tag{A.97}$$

$$\left(\delta F^{\rho}\right)^{\frac{1}{\rho-1}} = A - \frac{1}{2}\alpha x^2\sigma_K^2 - \frac{\delta}{\rho}\left[\left(\frac{F}{\delta}\right)^{\frac{\rho}{\rho-1}} - 1\right] \tag{A.98}$$

$$\delta^{\frac{1}{\rho-1}} F^{\frac{\rho}{\rho-1}}\left(1 - \frac{1}{\rho}\right) = A - \frac{1}{2}\alpha x^2\sigma_K^2 + \frac{\delta}{\rho}. \tag{A.99}$$

Hence,

$$F = \left[\left(\left(1-x\right)A_0 + xA_1 - \frac{1}{2}\alpha x^2\sigma_K^2 + \frac{\delta}{\rho}\right)\frac{\rho}{\rho-1}\delta^{-\frac{1}{\rho-1}}\right]^{\frac{\rho-1}{\rho}}, \tag{A.100}$$

or for EIS=1 case,

$$F = exp\left(\frac{1}{\delta}\left[\delta ln\delta - \frac{1}{2}\alpha x^2\sigma_K^2 - \delta + \left(1-x\right)A_0 + xA_1\right]\right) \tag{A.101}$$

$$c = \delta \tag{A.102}$$

$$i = A - \delta. \tag{A.103}$$

Maximize either of these wrt $x$:

$$x^* = min\left(\frac{A_1 - A_0}{\alpha\sigma_K^2}, 1\right). \tag{A.104}$$

Marginal $q$ in such economy is constant and equal 1 (on aggregate and for each individual technology). All the risk is thus cash flow risk, $\frac{1}{dt}\mathbb{E}\left[dR\right] = (A-i)+i = A$, $\frac{1}{dt}\mathbb{E}\left[dR_0\right] = A_0$, $\frac{1}{dt}\mathbb{E}\left[dR_1\right] = A_1$, $rx_M = rx_1 = \alpha x^*\sigma_K^2$, $rx_0 = 0$, and $r = A - \alpha x^*\sigma_K^2$.

### A.5.5 Two sector model, $EIS = 1$, log adjustment costs, time-varying $\alpha$

**SDF**

$$f = \delta J\left(lnC - lnJ\right) \tag{A.105}$$

$$f_C = \delta J\frac{1}{C} = \delta F\frac{1}{c} \tag{A.106}$$

$$dlnf_C = dlnF - dlnc \tag{A.107}$$

$$dlnJ = dlnK + dlnF \tag{A.108}$$

$$-sd\left(\frac{d\Lambda}{\Lambda}\right) = dlnc - dlnF + \alpha\left(dlnK + dlnF\right) \tag{A.109}$$

$$= dlnc + \alpha dlnK + \left(\alpha - 1\right)dlnF. \tag{A.110}$$

## A.6 Proofs

*Proof of Theorem 2.2.* In section A.2.1 I show simple manipulations of Duffie and Epstein (1992a,b) formulas to get risk prices for the case when the risk aversion $\alpha$ is constant. In section A.2.2 I derive the expression for market prices of risk in the general case when $\alpha$ is time-varying. □

*Proof of Theorem A.3.* I look for a first-order expansion of $F\left(x,\alpha;\epsilon\right)$ as a power series in $\epsilon$:

$$F\left(x,\alpha;\epsilon\right) = F\left(x,\alpha;0\right) + F_\epsilon\left(x,\alpha;0\right)\epsilon + o\left(\epsilon^2\right), \tag{A.111}$$

where $F\left(x,\alpha;0\right)$ gives the value of $F$ at $\epsilon = 0$ and $F_\epsilon\left(x,\alpha;\epsilon\right)$ gives the derivative at $\epsilon = 0$. To compute the first-order functional perturbation of Eq. A.69, I first evaluate the PDE in Eq. A.72 at $\epsilon = 0$ to find $F\left(x,\alpha;0\right) = \delta\left(A_0+\xi\right)^{1+\frac{\xi}{\delta}}\left(\delta+\xi\right)^{-1-\frac{\xi}{\delta}}$. Next, I differentiate the PDE with respect to $\epsilon$ and drop all terms multiplying $\epsilon$. I do so because these terms will always drop out in computations of all derivatives $\left.\frac{\partial^{k+1}F}{\partial\epsilon\partial x^k}\right|_{\epsilon=0}$ and $\left.\frac{\partial^{k+1}F}{\partial\epsilon\partial\alpha^k}\right|_{\epsilon=0}$ for any $k>0$ and evaluating at $\epsilon=0$. Next, all $\frac{\partial^{k+n}F}{\partial x^n\partial\alpha^k}=0$ for any $k$ and $n$ such that $k+n>0$. The resulting expression is:

$$-\bar{A}(\delta+\xi)x-\frac{1}{2}\alpha x^2\varsigma_K^2-\delta\frac{F_\epsilon}{F}+\phi(\bar{\alpha}-\alpha)\frac{F_{\epsilon x}}{F}=0. \tag{A.112}$$

Solving for $F_\epsilon$ and choosing a family of solutions in the space of real numbers delivers the expression for $F(x,\alpha;\epsilon)$ in Theorem A.3. First order Taylor expansion of $\omega_x(x,\alpha;\epsilon)\equiv\frac{F_x}{F}$ and $\omega_\alpha(x,\alpha;\epsilon)\equiv\frac{F_\alpha}{F}$ around $\epsilon=0$ gives the expressions for components of risk prices. □

*Proof of Theorem A.4.* Eq. A.43 gives expressions of marginal $q$'s as function of $\omega_x(x,\alpha)$ which we have already calculated in Theorem A.3. Performing Taylor expansion of these expressions around $\epsilon=0$ followed by a similar procedure discussed in the proof of Theorem A.3 to find $q(x,\alpha;0)$ and $q_\epsilon(x,\alpha;0)$ delivers the expressions in the theorem. Loadings $l_{n,x}(x,\alpha;\epsilon)\equiv\frac{q_{n,x}}{q_n}$, $l_{n,\alpha}(x,\alpha;\epsilon)\equiv\frac{q_{n,\alpha}}{q_n}$ are further calculated by Taylor-expanding the resulting expressions one more time. □

*Proof of Proposition A.2.* Unexpected returns of the low-risk technology in response to an orthogonal capital shock are given by (up to the first order in $\epsilon$):

$$dR_0-\mathbb{E}dR_0 \quad = \quad l_{0,x}x(1-x)\boldsymbol{\sigma_K}d\boldsymbol{Z} \tag{A.113}$$

$$\underset{\epsilon\to 0}{\simeq} \quad \left\{-\frac{\xi}{\delta}\bar{A}+2x(\zeta_0+\zeta_\alpha\alpha)\varsigma_K^2\right\}x(1-x)\sqrt{\epsilon}\boldsymbol{\sigma_K}d\boldsymbol{Z}. \tag{A.114}$$

According to Theorem A.2, for $x>x^*$ sufficiently close to $x^*$, $\omega_x=\frac{\delta+\xi}{\delta}\bar{A}-(\zeta_0+\zeta_\alpha\alpha)x\sigma_K^2<0$ and $\omega_x=0$ at $x=x^*$. Around the first order expansion in $\epsilon$, Theorem A.3 can be used to find that $F_{xx}=-(\zeta_0+\zeta_\alpha\alpha)\sigma_K^2<0$ does not change sign and the above result therefore holds globally on $x\in[x^*,\,1]$ (up to the first-order expansion in $\epsilon$). Plugging this in the formula above implies that the unexpected return on low-risk technology is positive when $x>x^*$.

At $x^*$, $(\zeta_0+\zeta_\alpha\alpha)x^*\varsigma_K^2=\frac{\delta+\xi}{\delta}\bar{A}$ producing

$$\mathcal{L}\left(dR_0-\mathbb{E}dR_0\Big|_{x=x^*}\right)\underset{\epsilon\to 0}{\simeq}\left(2+\frac{\xi}{\delta}\bar{A}\right)x^*(1-x^*)\varsigma_K, \tag{A.115}$$

which is positive. Evaluating the unexpected return at $x=k\,x^*$ for any constant $k<1$ gives

$$\mathcal{L}\left(dR_0-\mathbb{E}dR_0\Big|_{x=kx^*}\right)\underset{\epsilon\to 0}{\simeq}\left[-\frac{\xi}{\delta}\bar{A}+2k\left(1+\frac{\xi}{\delta}\right)\bar{A}\right]x^*(1-x^*)\varsigma_K, \tag{A.116}$$

which is positive when $k>\frac{1}{2}\frac{\xi}{\xi+\delta}$.

Therefore on the range $x\in[\frac{1}{2}\frac{\xi}{\xi+\delta}x^*,\,1]$ unexpected returns on the low-risk technology respond positively to an orthogonal capital shock.

Unexpected returns of the high-risk technology in response to an orthogonal capital shock are given by (up to the first order in $\epsilon$):

$$dR_1-\mathbb{E}dR_1 \quad = \quad [1+l_{1,x}x(1-x)]\boldsymbol{\sigma_K}d\boldsymbol{Z} \tag{A.117}$$

$$\underset{\epsilon\to 0}{\simeq} \quad \left\{1-\frac{\xi}{\delta}\bar{A}x(1-x)+(2x-1)(\zeta_0+\zeta_\alpha\alpha)x(1-x)\varsigma_K^2\right\}\sqrt{\epsilon}\boldsymbol{\sigma_K}d\boldsymbol{Z}. \tag{A.118}$$

I assume that the quantity in curly brackets is positive for any value of state variables (make it a standalone assumption and solve for required parameter values!). The assumption is fairly innocuous and holds for a wide range of plausible calibrations. It rules out the case when stock prices rise on a negative capital shock. With this assumption, returns on the high-risk technology react positively to an orthogonal capital shock. □

*Proof of Proposition A.3.* Consider the case of uncorrelated shocks, $\lambda=0$. It is sufficient to show that the proposition holds in this case. Unexpected returns of the low-risk technology are given by (up to the first order in $\epsilon$)

$$dR_0-\mathbb{E}dR_0 \quad = \quad [l_{0,x}x(1-x)+l_{0,\alpha}\alpha\lambda]\boldsymbol{\sigma_K}d\boldsymbol{Z}+l_{0,\alpha}\boldsymbol{\sigma_\alpha}d\boldsymbol{Z}, \tag{A.119}$$

and unexpected returns of the high-risk technology by

$$dR_1 - \mathbb{E}dR_1 = [1 + l_{1,x}x(1-x) + l_{1,\alpha}\alpha\lambda]\,\boldsymbol{\sigma_K} d\boldsymbol{Z} + l_{1,\alpha}\boldsymbol{\sigma_\alpha} d\boldsymbol{Z}. \tag{A.120}$$

When $\lambda = 0$ and up to a first order expansion in $\epsilon$, covariance of two returns is given by

$$cov_t(dR_0,\, dR_1) = l_{0,x}x(1-x)[1 + l_{1,x}x(1-x)]\,\varsigma_K^2 + l_{0,\alpha}l_{1,\alpha}\alpha^2\varsigma_\alpha^2. \tag{A.121}$$

The last term is negative and quadratic in $\alpha$. The first term is positive for $x \in [\frac{1}{2}\frac{\xi}{\xi+\delta}x^*,\, 1]$ and consist of two terms linear in $\alpha$ and one quadratic in $\alpha$, $2x(2x-1)(\zeta_0 + \zeta_\alpha\alpha)^2 x^2(1-x)^2\varsigma_K^6$. The quadratic term is decreasing in $\alpha$ for $x < \frac{1}{2}$. For high levels of $\alpha$ quadratic terms dominate and thus the covariance between two returns becomes more negative. It is therefore possible to calibrate a model in such a way that the covariance is positive for low levels of $\alpha$ (when $-l_{0,\alpha}l_{1,\alpha}\alpha^2\varsigma_\alpha^2$ is relatively small and $l_{0,x}x(1-x)[1 + l_{1,x}x(1-x)]\,\varsigma_K^2$ is positive and dominates) and negative when $\alpha$ becomes high, as quadratic terms start dominating and $l_{0,x}x(1-x)[1 + l_{1,x}x(1-x)]\,\varsigma_K^2$ becomes small.

□

*Proof of Proposition A.4.* When $\lambda = 0$, the unexpected returns of the high-risk technology are given by

$$dR_1 - \mathbb{E}dR_1 \underset{\epsilon\to 0}{\simeq} \left\{1 - \frac{\xi}{\delta}\bar{A}x(1-x) + (2x-1)(\zeta_0 + \zeta_\alpha\alpha)\,x(1-x)\,\sigma_K^2\right\}\sqrt{\epsilon}\sigma_K d\boldsymbol{Z} + l_{1,\alpha}\alpha\sigma_\alpha d\boldsymbol{Z}. \tag{A.122}$$

The loading on capital shock therefore increases in $\alpha$ when $x < \frac{1}{2}$. The loading on risk aversion risk decreases in $\alpha$ since $l_{1,\alpha}$ is negative and decreasing in $\alpha$.

The loading of the SDF on shocks is given by

$$\mathcal{L}\left(-\frac{d\Lambda}{\Lambda}\right) = \bar{A}(1-x)\,x\sigma_K + \alpha x\sigma_K + (\alpha-1)\,\omega_x(x,\alpha)\,x(1-x)\,\boldsymbol{\sigma_K} + (\alpha-1)\,\omega_\alpha(x,\alpha)\,\alpha\boldsymbol{\sigma_\alpha}. \tag{A.123}$$

The price of the $\alpha$-shock risk is therefore negative and decreasing in $\alpha$ (for $\alpha > 1$), since $\omega_\alpha(x,\alpha)$ is negative. Since both the loading and the price of risk aversion risk are negative and decreasing in $\alpha$, the risk premium due to the risk aversion always increases in the level of $\alpha$. The price of the capital risk is given by

$$\left[\bar{A}(1-x) + \alpha + (\alpha-1)\,\omega_x(x,\alpha)(1-x) + (\alpha-1)\,\omega_\alpha(x,\alpha)\,\alpha\lambda\right]x\sigma_K \underset{\epsilon\to 0}{\simeq} \tag{A.124}$$

$$\left[\bar{A}(1-x) + 1 + (\alpha-1)\left(1 + \left[\zeta_A\bar{A} - (\zeta_0 + \zeta_\alpha\alpha)\,x\sigma_K^2\right](1-x)\right)\right]x\sigma_K. \tag{A.125}$$

By expressing $x$ as $x = k\,x^*$, the last term can be rewritten as $(\alpha-1)\left(1 + (1-k)\frac{\delta+\xi}{\xi}\bar{A}(1-x)\right)$. When $k < 1 + \frac{\xi}{\xi+\delta}\frac{1}{\bar{A}}$, the price of capital risk is always positive and increasing in $\alpha$ and thus the risk premium due to capital risk is also increasing in $\alpha$. The overall risk premium then must be monotonically increasing in $\alpha$.

□

*Proof of Theorem 3.1.* Plug $\boldsymbol{\sigma_J}$ inside Eq. A.6,

$$0 = \boldsymbol{\lambda} + \left(\frac{J_{WW}}{J_W} - \alpha\frac{J_W}{J}\right)W\boldsymbol{\sigma_R}\boldsymbol{\sigma_R}'\boldsymbol{\theta} + \left(\frac{J_{W\boldsymbol{X}}}{J_W} - \alpha\frac{J_{\boldsymbol{X}}}{J}\right)\boldsymbol{\sigma_R}\boldsymbol{\sigma_X}'. \tag{A.126}$$

The value function is homogeneous in wealth, $J = W \times G(\boldsymbol{X})$. Then,

$$J_{WW} = 0 \tag{A.127}$$

$$\frac{WJ_W}{J} = 1 \tag{A.128}$$

$$\frac{J_{W\boldsymbol{X}}}{J_W} = \frac{J_{\boldsymbol{X}}}{J} = \frac{G'(\boldsymbol{X})}{G(\boldsymbol{X})}. \tag{A.129}$$

Plugging this in gives

$$\boldsymbol{\lambda} = \alpha \boldsymbol{\sigma_R} \boldsymbol{\sigma}_{\boldsymbol{R}}^{'} \boldsymbol{\theta} + (\alpha - 1) \frac{J_{\boldsymbol{X}}}{J} \boldsymbol{\sigma_X} \boldsymbol{\sigma}_{\boldsymbol{R}}^{'} \tag{A.130}$$

$$\boldsymbol{\mu_R} - r = \alpha \times cov\left(dR, dR_{TW}\right) + \eta \times cov\left(dR, d\boldsymbol{X}\right), \tag{A.131}$$

where $\eta = (\alpha - 1) \frac{J_{\boldsymbol{X}}}{J}$, $R_{TW}$ is return on the total wealth portfolio, $d\boldsymbol{X} = (\alpha, x)^{'}$. □

## A.7 Computational Details

### A.7.1 Solution Method

The model is solved numerically by finding a solution to the system of equations in Theorem 2.1 using high-order projection methods. I parameterize the value function and two investment functions as a complete product of $20^{th}$-order Chebyshev polynomials in two state variables, $x$ and $\alpha$. Once the approximation to the value and investment functions are found, I use them to calculate the aggregates and prices in the model. section A.7.1 provides further details.

I find a numerical solution to the system of equations in Theorem 2.1 using high-order projection methods. I parameterize the value function and two investment functions as a complete product of $20^{th}$ order Chebyshev polynomials in two state variables, $x$ and $\alpha$. Next, I evaluate the system of equations in Theorem 2.1 at $30 \times 30$ points on the state space. Points are chosen as Chebyshev's zeroes. I then search for coefficients of three policy functions to minimize the $L_1$-norm of PDE errors. In practice, the algorithm is iterative. I start by fitting low order polynomials on the grid of $30 \times 30$ Chebyshev zeroes and iteratively increase the order of the fit until the desired precision is reached. I impose boundary conditions as given in section A.3.3. The resulting fit is the global solution on the entire state space.

The problem is therefore formulated as a sequence of standard constrained optimization problems with thousands of constraints (one at each grid node) and thousands of unknowns (Chebyshev coefficients). I use the GAMS modeling language together with CONOPT and SNOPT non-linear constrained optimizers to find a solution. Once the solution to the optimization problem is found, I import the results in Matlab to perform simulations and report the results.

Once the approximation to the value and investment functions is found, I can use them to calculate the aggregates and prices in the model. Refer to section A.3 for further details.

### A.7.2 Small-noise expansions

I use Mathematica to calculate analytical derivatives required by the small-noise expansions. First order expansions are straightforward to derive as described in section A.6. Expansions of higher orders can be computed in a similar fashion. I computed expansions up to a third order (analytically) to verify that the directions of the main forces in the first order expansions are not overturn by higher orders. I computed expansions up to the $10^{th}$ order numerically (using high precision arithmetic) as well.

### A.7.3 Projections

I use high order projection methods described in Judd (1998) to solve the general model. In particular, I parametrize the value function and two investment functions as a complete product of $20^{th}$ order Chebyshev polynomials in two state variables, $x$ and $\alpha$. Next, I evaluate the system of equations in Theorem 2.1 at $30 \times 30$ points on the state space. Points are chosen as Chebyshev's zeroes. I then search for coefficients of three policy functions to minimize the $L_1$ norm of PDE errors. In practice the algorithm is iterative. I start by fitting low order polynomials on the grid of $30 \times 30$ Chebyshev zeroes and iteratively increase the order of the fit until the desired precision is reached. I impose boundary conditions as given in section A.3.3. The resulting fit is the global solution on the entire state space.

The problem is therefore formulated as a sequence of standard constrained optimization problems with thousands of constraints (one at each grid node) and thousands of unknowns (Chebyshev coefficients). I use the GAMS modeling language together with CONOPT and SNOPT non-linear constrained optimizers to find a solution. Once the solution to the optimization problem is found, I import the results in Matlab to perform 1,000,000 simulations and report the results. Most time-sensitive parts of the code for simulations and impulse responses have been programmed in C++ for faster execution.

#### A.7.4 Impulse responses

I compute impulse responses by shocking the economy at its steady state (unconditional means of state variables) and performing Monte-Carlo simulations for the following $20 \times 12$ months. I simulate 10,000,000 Monte-Carlo trajectories for each shock and calculate the means of realizations of moments of interest. The high number of simulations is necessary due to high volatility of the SDF and persistence in the state variables.

#### A.7.5 Term structure

To compute the term structure I calculate conditional expectations of the SDF at each horizon iteratively starting from the end. This requires fitting the price of a bond as function of state variables at each iteration. I do so by evaluating the conditional expectation at each node (Chebyshev zeroes) and then fitting a smooth function of two state variables to these points. The function is constructed as a complete product of two 10 degree Chebyshev polynomials. The fitting requires a search for the coefficients of this function (55 coefficients). I call GAMS within my Matlab code to compute each fit. Finally, I use Gauss-Hermite quadrature to compute the required integrals (conditional expectations).

#### A.7.6 Calibration: the Neural Network

Section 3.2.1 provides the summary of how the model is calibrated. In this section I detail how steps 4 and 5 in this procedure are implemented.

The task of the neural network is to create a functional mapping from the space of parameters into the space of model's implied moments. Considering that the mapping is highly non-linear and high-dimensional, a neural network is an ideal candidate to approximate it.

The process starts with generating the "training data" — solutions to the model for various choices of admissible parameters (inputs) and calculating the moments implied by each model (outputs). The inputs and outputs to the model are standardized based on the observed distributions in the training data. I parameterize the network to contain four layers of neurons, with 32 neurons in each layer. I use *tanh* as an activation function and employ batch normalization in each layer of the network. The final layer uses a linear activation function.

I use Google's TensorFlow package to implement and estimate the model. The model is estimated by back-propagation – picking random weights, differentiating the network through with respect to these parameters, and continuously adjusting the value of the parameters using gradient descent in order to minimize the least square distance between the network's predicted and model-implied moments.

Once the model is fitted, it delivers a non-linear approximator from the space of parameters into the space of moments. This functional approximator is fully differentiable with respect to the inputs (model's parameters), with derivatives readily available via backpropagation and the chain rule. I, therefore, differentiate this approximator with respect to inputs (model's parameters) and adjust these parameters (using gradient descent) in a way that minimizes the least squared distance between the network's predicted moments and their empirical targets. The final set of the model's parameter values is recorded.

In practice, because neural networks are randomly initialized, the last two steps are repeated 100 times to estimate an ensemble of neural networks with their predictions subsequently averaged out.

Finally, the model is re-estimated using neural network's predicted parameter values.

#### A.7.7 Additional Model-Implied Distributions and Results

Figures A.1 and A.2 show histograms of variables conditional on the share of high-risk capital, $x$.

# B Additional Empirical Results

## B.1 Return predictability regressions using the level of correlation

To partially alleviate issues with non-stationarity in Bond-Stock correlation, share of risky capital, and inflation dynamics in the early part of the sample, I focus on TIPS returns and their respective sample from 1999 till 2020 in Table B.1. The table uses the level of bond-stock correlation as a predictor, $Corr$. The table shows that coefficients on $Corr$ are generally smaller in magnitude than the ones on $\Delta Corr$ in Table 6 in the paper, but are qualitatively similar. There isn't enough statistical evidence to reject the null hypothesis of zero, however. Table B.2 looks at the predictability of the long short portfolios of equities based on definitions of low- and high-risk industries in Section 4.2.2. The table reports results at multiple horizons $h = 1..5$ years (predict $h$-year future returns, $r_{t\to t+h}$, using the

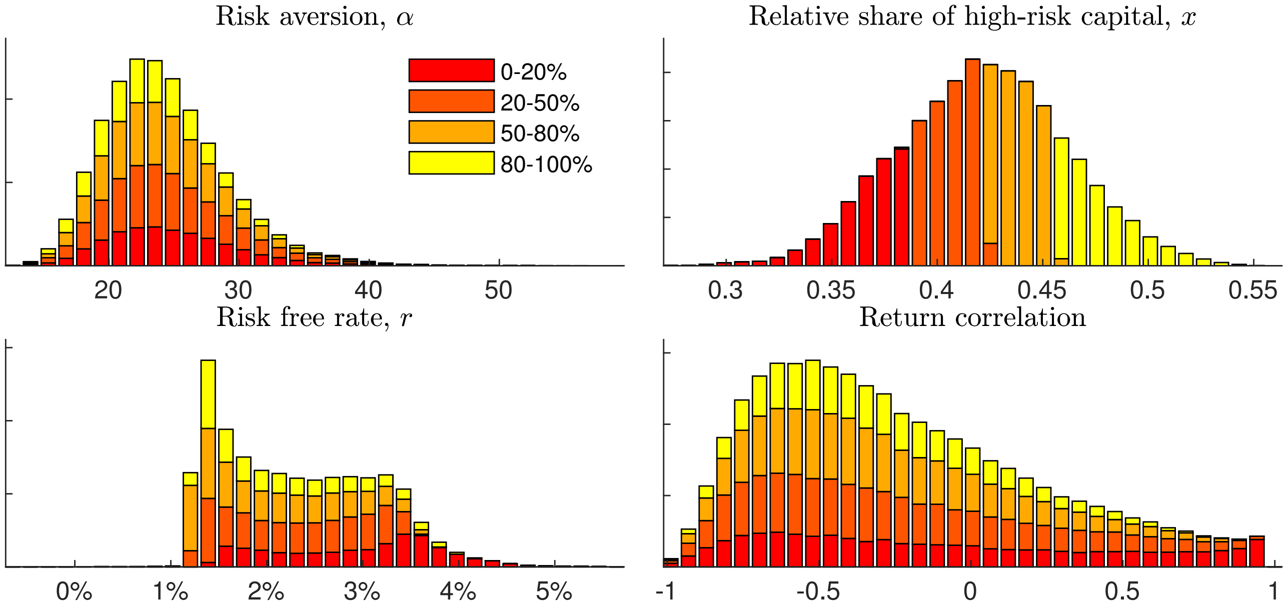



Figure A.1: **Conditional distributions of state variables, interest rate, and Bond-Stock correlations (conditional on $x$).** Each color within a bar shows a fraction of simulated observations with the share of high-risk capital, $x$, within a given percentile.

level of bond-stock correlation, $Corr_t$, in the full sample). The table shows the lack of statistical evidence at $h = 1$, but strong evidence against the null of zero at $h = 5$.

## B.2 Return predictability regressions using changes of correlation across multiple horizons

As additional robustness evidence I conduct additional tests to show this relationship holds true in the data at multiple predictive horizons, consistent with the impulse response intuition discussed earlier. Table B.3 shows that the coefficient of a long stocks and short bonds portfolio $h$-year returns, $r_{t\to t+h}$, on $h$-year changes in correlation, $\Delta Corr_{t-h\to t}$, becomes more economically and statistically significant as the predictability horizon $h$ increases. Table B.4 shows similar evidence for the long-short portfolio based on definitions of low- and high-risk industries in Section 4.4.2.

Table B.5: Yield Curve and Bond-Stock Correlations (3-year horizon)

The table shows estimates of a regression of 3-year changes in level and slope of the yield curve onto the changes in bond-stock correlations. Nominal treasuries in the first two columns, TIPS in the last two columns. $t$-statistics in parantheses use HAC standard errors with a small-sample correction.

| | Nominal rates | | TIPS | |
|---|---|---|---|---|
| | ΔLevel | ΔSlope | ΔLevel | ΔSlope |
| ΔCorr | 0.024 | -0.033 | 0.023 | -0.310 |
| | (5.23) | (-2.54) | (2.39) | (-2.75) |
| Obs. | 11,234 | 11,234 | 4,477 | 4,477 |
| $R^2$ | 0.13 | 0.19 | 0.21 | 0.24 |

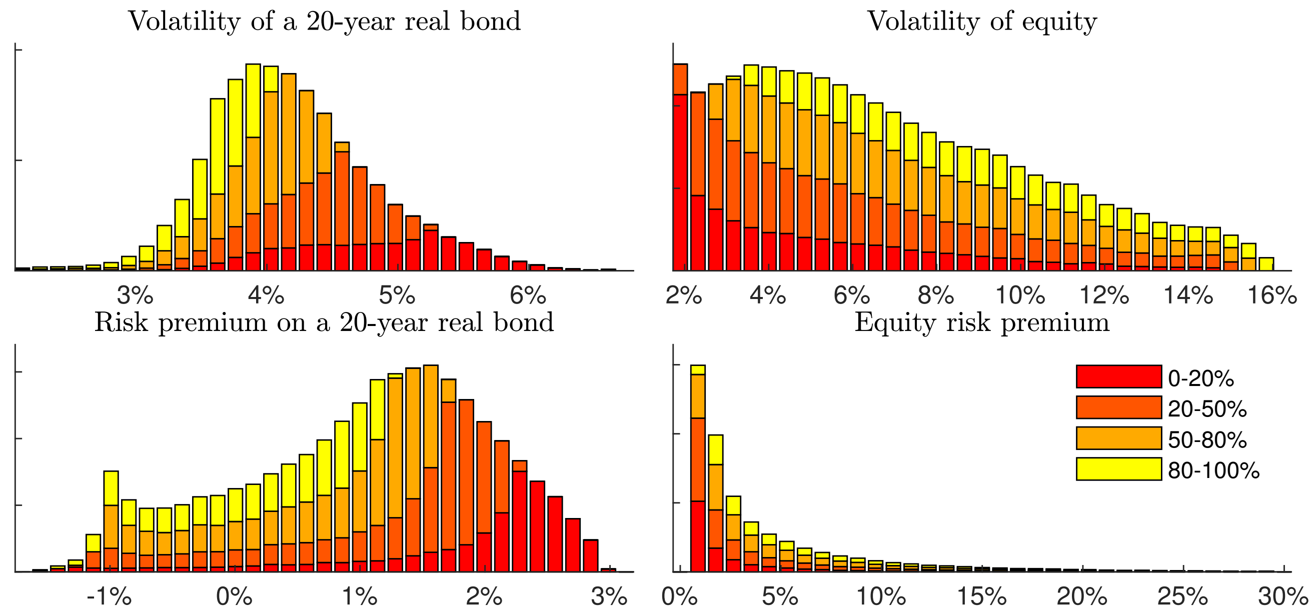



Figure A.2: **Conditional distributions of financial variables (conditional on $x$).** Each color within a bar shows a fraction of simulated observations with the share of high-risk capital, $x$, within a given percentile.

Table B.1: Dynamics of Bond and Stock risk premia: TIPS.

The table reports predictability of term premia on a 10-year TIPS ($r_b$), equity index premia ($r_s$), and their difference ($r_s - r_b$) with the level of Bond-Stock correlation. The last three columns include controls for equity index volatility and log equity Price-Dividend ratio. $t$-statistics in parantheses use HAC standard errors with a small-sample correction.

| | $r_b$ | $r_s$ | $r_s - r_b$ | $r_b$ | $r_s$ | $r_s - r_b$ |
|---|---|---|---|---|---|---|
| Corr | 0.03 | -0.01 | -0.05 | 0.04 | -0.04 | -0.08 |
| | (0.59) | (-0.06) | (-0.25) | (0.76) | (-0.14) | (-0.34) |
| Vol | - | - | - | -0.56 | -7.39 | -6.38 |
| | - | - | - | (-0.75) | (-3.47) | (-3.43) |
| pd | - | - | - | -0.06 | -0.20 | -0.13 |
| | - | - | - | (-1.27) | (-1.03) | (-0.65) |
| Obs. | 4,981 | 4,981 | 4,981 | 4,981 | 4,981 | 4,981 |
| $R^2$ | 0.01 | 0.00 | 0.00 | 0.04 | 0.07 | 0.04 |

Table B.6: Yield Curve and Bond-Stock Correlations (with Vol.).

The table shows estimates of a regression of annual changes in the level and slope of the yield curve onto the changes in bond-stock correlations, controlling for changes in realized stock volatility. Nominal treasuries in the first two columns, TIPS in the last two columns. $t$-statistics in parantheses use HAC standard errors with a small-sample correction.

| | Nominal rates | | TIPS | |
|---|---|---|---|---|
| | ΔLevel | ΔSlope | ΔLevel | ΔSlope |
| ΔCorr | 0.018 | -0.013 | 0.015 | -0.042 |
| | (4.48) | (-1.90) | (3.11) | (-0.38) |
| ΔVol | -0.093 | 0.008 | 0.242 | -5.257 |
| | (-0.69) | (0.05) | (1.38) | (2.21) |
| Obs. | 11,738 | 11,738 | 4,981 | 4,981 |
| $R^2$ | 0.12 | 0.05 | 0.26 | 0.15 |

Table B.2: Long-Short Equity Predictability with the level of Bond-Stock Correlation.

The table reports coefficient estimates and the $R^2$ of a regression of future $h$-year returns on the long-short portfolio based on the definitions of low- and high-risk industries in Section 4.4.2, $r_{t\to t+h}$, onto the level of Bond-Stock correlation, $Corr_t$. The columns correspond to values of $h = 1..5$. $t$-statistics in parantheses use HAC standard errors with a small-sample correction.

| $h$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Corr | 0.01 | -0.00 | -0.05 | -0.11 | -0.25 |
| | (0.17) | (-0.05) | (-0.74) | (-1.57) | (-2.86) |
| Obs. | 7,604 | 7,107 | 6,618 | 6,123 | 5,626 |
| $R^2$ | 0.00 | 0.00 | 0.02 | 0.05 | 0.13 |

Table B.3: Predictability of Stock-minus-Bond returns with changes in Bond-Stock correlation.

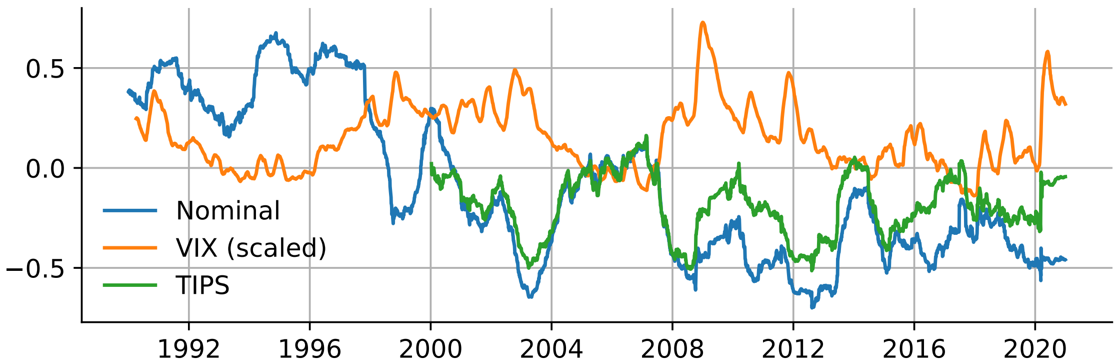
The table reports $h$-year return predictability on the long-short portfolio of the equity market index and a 10-year bond, $r_{t\to t+h}$, with $h$-year differences in Bond-Stock correlation, $\Delta Corr_{t-h\to t}$. The columns correspond to values of $h = 1..5$. $t$-statistics in parantheses use HAC standard errors with a small-sample correction.

| $h$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| ΔCorr | -0.22 | -0.18 | -0.38 | -0.46 | -0.44 |
| | (-3.76) | (-1.90) | (-2.59) | (-3.84) | (-6.03) |
| Vol | -5.55 | 4.86 | 1.64 | -9.42 | -11.24 |
| | (-3.26) | (1.01) | (0.51) | (-1.36) | (-1.67) |
| pd | -0.01 | -0.04 | -0.08 | -0.09 | -0.19 |
| | (-0.11) | (-0.43) | (-1.24) | (-2.00) | (-1.88) |
| Obs. | 11,739 | 11,235 | 10,731 | 10,227 | 9,723 |
| $R^2$ | 0.08 | 0.06 | 0.16 | 0.15 | 0.12 |

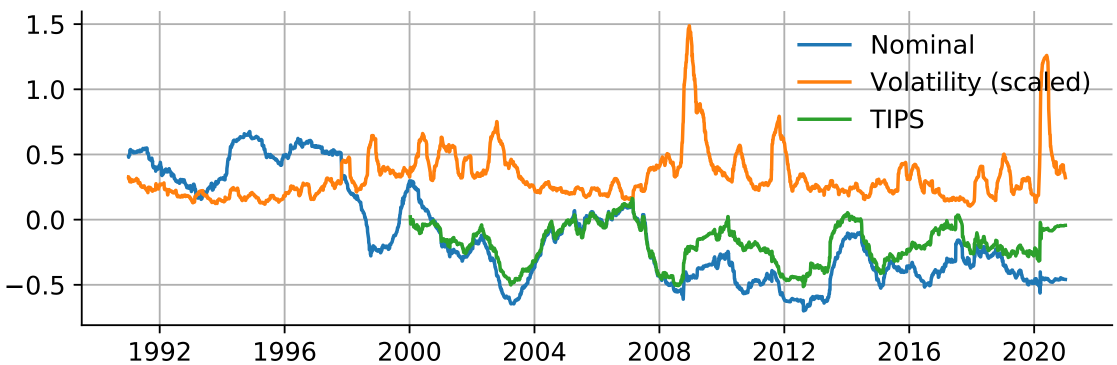
Table B.4: Long-Short Equity Portfolio Predictability with changes in Bond-Stock Correlation.

The table reports coefficient estimates and the $R^2$ of a regression of future $h$-year returns on the long-short portfolio based on definitions of low- and high-risk industries in Section 4.4.2, $r_{t\to t+h}$, onto the $h$-year changes in Bond-Stock correlation, $\Delta Corr_{t-h\to t}$. The columns correspond to values of $h = 1..5$. $t$-statistics in parantheses use HAC standard errors with a small-sample correction.

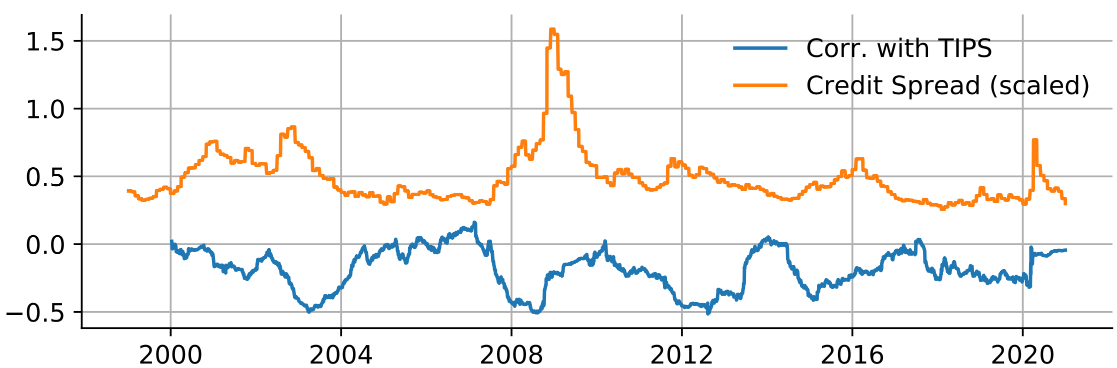
| $h$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| ΔCorr | -0.07 | -0.08 | -0.15 | -0.22 | -0.24 |
| | (-1.51) | (-1.67) | (-2.68) | (-5.99) | (-9.45) |
| Obs. | 7,604 | 7,107 | 6,618 | 6,123 | 5,626 |
| $R^2$ | 0.02 | 0.03 | 0.15 | 0.31 | 0.34 |




(a) VIX




(b) 3-month realized volatility




(c) GZ Credit spread

Figure B.1: **Bond-stock correlation and empirical proxies for risk aversion.** Time series of rescaled values of smoothed VIX (3-month moving-average; panel (a)), realized 3-month volatility (panel (b)), or GZ credit spreads (panel (c)), shown in orange, and rolling 1-year correlations between daily stock excess returns and 10-year bond excess returns (nominal and real, in blue and green, respectively).

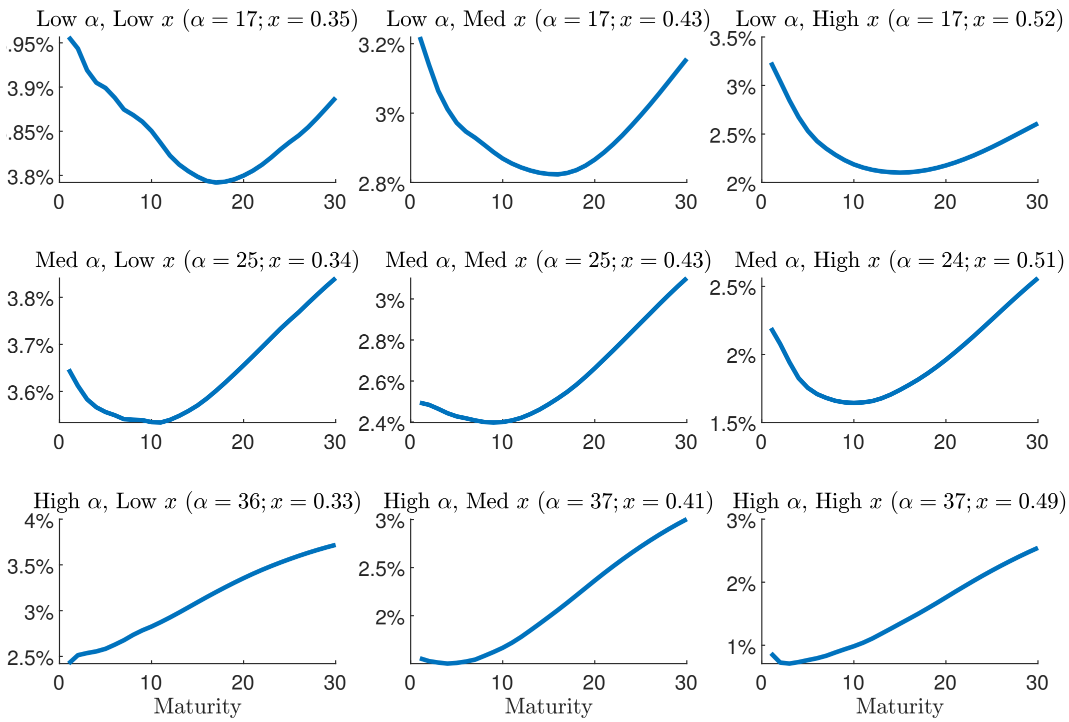



Figure B.2: **Average conditional yield curves.** The curves are conditional on the level of risk aversion and the share of risky capital. The levels of $\alpha$ and $x$ can take one of the three values: Low (5th percentile), High (95th percentile), and Medium.

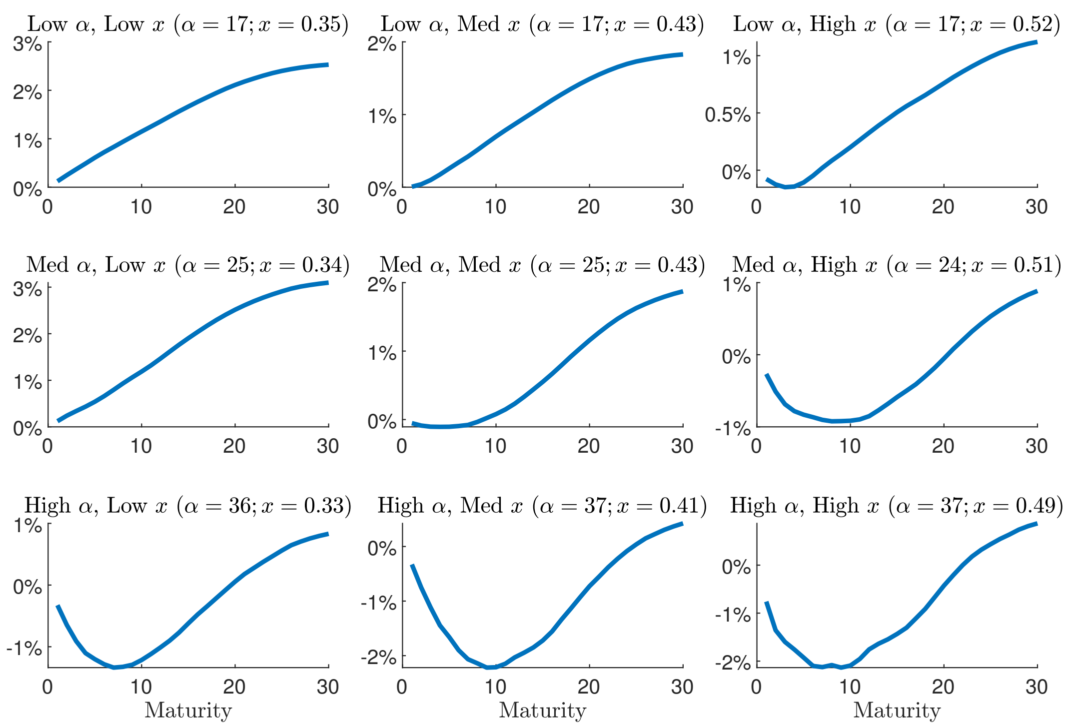



Figure B.3: **Average conditional term premia curves.** The curves are conditional on the level of risk aversion and the share of risky capital. The levels of $\alpha$ and $x$ can take one of the three values: Low (5th percentile), High (95th percentile), and Medium.

# References


Anderson, E., L. Hansen, and T. Sargent (2012). Small noise methods for risk-sensitive/robust economies. *Journal of Economic Dynamics and Control 36*(4), 468.

Duffie, D. and L. Epstein (1992a). Asset pricing with stochastic differential utility. *Review of Financial Studies 5*(3), 411–436.

Duffie, D. and L. Epstein (1992b). Stochastic differential utility. *Econometrica: Journal of the Econometric Society*, 353–394.

Fleming, W. (1971). Stochastic control for small noise intensities. *SIAM Journal on Control 9*, 473.

Fleming, W. and J. Yang (1994). Numerical methods for infinite horizon risk sensitive stochastic control. In *Decision and Control, 1994., Proceedings of the 33rd IEEE Conference on*, Volume 2, pp. 1945–1949. IEEE.

James, M. and M. Campi (1996). Nonlinear discrete-time risk-sensitive optimal control. *International Journal of Robust and Nonlinear Control 6*, 1–19.

Judd, K. (1998). *Numerical methods in economics.* The MIT press.

Kogan, L. and R. Uppal (2001). Risk aversion and optimal portfolio policies in partial and general equilibrium economies.