Price Dynamics in Partially Segmented Markets *

Robin M. Greenwood       Samuel G. Hanson       Gordon Y. Liao
Harvard University and NBER  Harvard University and NBER  Harvard University

Preliminary first draft: August 29, 2014

Abstract

Even when the forces of arbitrage prevail ultimately, financial markets may be segmented in the short run due to the role of market specialists. We develop a model of securities prices and yield curve dynamics in which capital moves quickly within a market, but more slowly between markets for different assets. We show how supply shocks in one market are transmitted over time, in both prices and quantities, to secondary markets. We discuss several applications, including the design and evaluation of government quantitative easing programs, and the role of corporate issuance in market integration.

*We thank Daniel Bergstresser, Jeremy Stein, and participants at Brandeis for useful feedback.
1 Introduction

In this paper, we develop a model of price and yield curve dynamics in which capital moves quickly within a market, but more slowly between markets for different assets. Our main objective is to understand how supply and demand shocks are reflected in prices and investor behavior in neighboring markets.

It has long been recognized that real world capital markets are segmented, meaning that risk is shared imperfectly across the boundaries of markets, and especially asset classes. Segmentation arises because of a variety of institutional, informational, and behavioral frictions that lead some investors to specialize. The impact of specialization is particularly visible when markets encounter shocks. Front-line specialists who can quickly absorb a shock have a limited risk appetite, leading to steep demand curves. When specialists are unable or unwilling to trade across markets, demand curves in one asset can appear disconnected from other assets.

While markets may be quite segmented at short horizons, in the long-run the forces of competition lead to integration. For example, Cochrane (2011) points out that “markets can be segmented until ‘deep pockets’ arrive, but they do arrive eventually.” “Eventually” may turn out to be a long time, however. Investment committees at pension funds and endowments typically reallocate capital annually; other investors may reallocate even more slowly. Duffie and Strulovici (2012) explain how differences in capital levels across markets provide incentives for reallocation of capital, and thus affect the speed of reallocation.

Empirically, there are numerous examples of capital moving slowly to dislocated markets. Examples include Froot and O’Connell (1999), who show that capital flows slowly in the catastrophe reinsurance industry following large losses, and Mitchell, Pedersen, and Pulvino (2008) who suggest that slow moving capital was responsible for a dislocation in the convertible bond market.

The fixed income markets provide a natural setting to explore the dynamics of slow moving capital. Assets in fixed income markets have a high degree of substitutability as the common exposure to the risk free rate constitutes a significant fraction of the overall risk in each market. The high degree of substitutability allows market participants to more readily integrate the market. In contrast, price movements in single-name equities are dominated by idiosyncratic cash-flow risks. A second more technical reason to study fixed income markets is that risk premia are naturally observed in yield space. The interest rate term structure and the spreads between treasury curve and other bond curves offer an intuitive way to decompose risk premia into different components.

We present a model that explicitly tackles the dynamics of security prices at different horizons across markets. We consider a setting with two risky coupon-paying bonds that trade in partially segmented markets. The bonds are partial substitutes, meaning that absent frictions, the prices of the two securities would be tightly linked by arbitrage. For concreteness, it may be useful to think of the primary market as being the US Treasury market, and the second market as being the corporate...
bond market. We consider a market structure in which two sets of front-line specialists can flexibly trade each type of bond in conjunction with the risk-free asset, but are unable to reallocate capital across the two bonds. Markets are partially integrated by generalist investors, who periodically reevaluate their portfolios and shift capital between bonds. This setup is similar to Gromb and Vayanos (2002), except that we constrain the cross-market arbitrageurs to be slow-moving, so as to be able to explore the dynamics of securities prices in the two markets.

Consider what happens when there is an unanticipated supply shock in one market, perhaps driven by the Federal Reserve buying some Treasury bonds. In the primary market (US Treasuries), primary market specialists, such as treasury bond dealers and specialized funds, immediately react by supplying the bonds out of their inventories. The excess demand drives the price higher and yield lower. Prices and yields in this market overreact, however, because the amount of capital that can accommodate the shock is limited to specialists in the short-run. The dynamics we describe are similar to those described in Duffie’s (2011) regarding asset prices when there is slow moving capital.

Our more novel results pertain to the behavior of prices and yields in the secondary market. Recall that this market has not received a direct shock, and is thus only indirectly affected through the behavior of the generalists. If the markets were fully integrated, securities in the secondary market would reprice immediately. However, generalists, such as pension funds, endowments and households, reallocate their portfolios slowly. Over time, these investors are able to digest the shock and reallocate part of their Treasury holdings in the primary market into corporate bonds holdings in the secondary market given the lower expected future returns on Treasury bonds after the price rally. The prices of the corporate bonds subsequently rise and the Treasury bond prices partially revert. Thus, the generalist investors act as a force of integration between the two segmented yet similar bond markets. The slow integration by generalists can result in overreaction of price and return in the primary market and under-reaction in the secondary market. The convergence in risk premia following a shock is never fully complete, however, because generalists command a risk premium for integrating the markets. In other words, even though capital is completely mobile in the long-run, the prices of risk across markets is not generally the same, because the generalists demand compensation for a form of noise-trader risk.

The results above depend critically on the fractions of specialists in each market, the number of time periods over which generalists rebalance portfolios, and the degree of substitutability across markets. Consider first the role played by the population sizes of generalists and specialists. When there are a small number of generalists, the primary market overreacts while the secondary market under-reacts in response to a supply shock. When there are no generalist market participants, the markets are completely segmented and the shock is not transmitted between markets. When there are many generalists, the markets are well-integrated and we have overreaction in both markets. Holding the fraction of generalists the same, increasing the fraction of specialists in the primary
market versus specialists in the secondary market results in less overreaction in the primary market with little impact on the secondary market. Now consider the role of generalists’ portfolio re-balancing time. With longer portfolio re-balancing time, or a smaller fraction of generalists re-balancing their portfolios in each period, the magnitudes of overreaction (under-reaction) in the primary (secondary) market are greater. Finally, consider the role of market-specific risk. When idiosyncratic risk in the secondary market is high, both short-run and long-run integration between the markets are harder to achieve, and thus the price impact is only significant in the market that is affected by the shock.

Our model also allows us to construct full zero-coupon rate curves to analyze the impact of supply shocks on the term structures of interest rates in the two markets. Consistent with empirical observations, we show that the yield curves for both markets steepen in presence of a positive supply shock and flatten in presence of a negative supply shock. This matches with the yield curve behaviors during QE and taper. The term-structure changes characterize the “reaching-for-yield” behavior of investors, buying bonds of longer maturities as overall risk premia decrease.

We also explore the dynamics of pre-announced supply shocks using our model. Prices in both the primary and secondary market react immediately to the announcement of a future, positive shock in the primary market. However, the short-run price changes exhibit substantial under-reaction in comparison with the long-run adjustments. The generalists begin a gradual adjustment towards the direction of the shock immediately following the announcement by exchanging secondary market holdings into primary market holdings. The specialists in the primary market first provide supply to the generalists by selling their holdings and then purchase the excess supply once the shock is delivered into the market. The result is a prolonged and gradual adjustment of prices starting from the time of announcement to after the delivery of the shock. Longer delays between the announcement and the delivery of the shock result in a more gradual adjustment process.

Besides being of theoretical interest, the question of how prices adjust across partially segmented markets is one of enormous practical importance. Consider the evaluation of quantitative easing policies conducted between 2009 and 2013 by the Federal Reserve, Bank of England, and more recently the European Central Bank. The favored methodology for evaluating the impact of supply shocks such as these has been the event study. In the typical event study, the researcher or policy maker attempts to measure the statistical significance of short-run price changes (or yield changes) following a policy announcement. The conclusions differ enormously. Focusing on short horizons after announcement dates, Krishnamurthy and Vissing-Jorgensen (2013) suggest that the effects of QE are largely isolated in the market in which the Federal Reserve is transacting in. This corresponds to a view of extreme market segmentation. Mamaysky (2014) suggests that if one expands the measurement window by a few days or weeks, however, the effects in secondary markets may be larger by an order of magnitude.

More generally, our model suggests that short-run price responses to announcements of supply
changes can be extremely misleading, especially so when we compare prices across different markets. We illustrate this idea more formally by analyzing the statistical power of event studies within our model. The horizon at which statistical power is maximized is often much shorter than the horizon at which the long-run price response is achieved. As a suggestive test of this idea, we revisit Mamaysky’s findings regarding the time horizon of the price impact of QE policies. Consistent with the model, we show that in a low rate volatility and low supply volatility environment characteristic of the QE period, the primary market exhibits large and statistically significant price responses to supply shocks while the secondary market shows meaningful yet statistically undetectable price changes.

Our findings are closely tied to two strands of research in finance. The idea that front-line arbitrageurs in securities markets are specialized traces back to Merton (1987) and Grossman and Miller (1988), and is a central part of the theory underlying limits-of-arbitrage (De Long et al 1990, Shleifer and Vishny 1997, Gromb and Vayanos 2002). More recently, a number of researchers have demonstrated locally downward sloping demand curves for financial assets, which would be puzzling if markets were fully integrated (Krishnamurthy et al (2007), Hanson (2014)).

Second, our paper is related to ongoing research on “slow-moving capital.” Here, our model is most closely related to Duffie (2010) and Duffie and Strulovici (2012), who also present a model of the movement of capital across two partially segmented markets. The focus in Duffie and Strulovici (2012), however, is on the endogenous determinants of capital mobility, which we take as exogenous. Our contribution is to explicitly characterize the dynamics of prices and yields.

2 Model

We develop the model in two steps. We first develop a discrete-time benchmark model of the term structure of interest rates, which is a discrete time analog of Vayanos and Vila (2009) and Greenwood and Vayanos (2014). This serves as a useful benchmark for the analysis that follows, because yields and expected returns can be easily decomposed into unconditional and conditional risk premia that depend on asset supply and the risk exposures of the securities in the market.

In the second step, we introduce a second market in which there is a security that is similar to the security in the first market, except that the security is exposed to an additional form of risk — cash flow risk — making it an imperfect substitute for the first risky asset. Here, we consider a market structure in which specialists can trade flexibly in their respective markets, but constrained generalist investors play the role of integrating the markets.

2.1 Single Asset Model

We start by developing a benchmark model of a single asset. The market contains a single coupon paying risky asset as well as a risk-free asset. The risk-free asset pays a coupon $r_t$ that is known in
the beginning of each period and follows an exogenous stochastic process

\[ r_{t+1} = \bar{r} + \rho_r (r_t - \bar{r}) + \varepsilon_{r,t+1} \tag{1} \]

where \( \text{Var}_t [\varepsilon_{r,t+1}] = \sigma^2_r \).

To simplify the exposition, we endow the market with a coupon paying perpetual bond. With positive interest rates, this allows us to measure net supply of the asset as a function of the duration. Let \( P^L_t \) be the price of the perpetual bond paying fixed coupon of \( C \) each period. The gross return on the risky asset is \( 1 + R^L_{t+1} = (P^L_{t+1} + C) / P^L_t \), where \( P^L_{t+1} \) is the price tomorrow. Using Campbell-Shiller (1988) log-linear approximation and defining \( \theta \equiv 1 / (1 + C) \), the one-period excess log return, \( r_{x,t+1}^L \), of the perpetual bond over the risk-free rate is given by

\[ r_{x,t+1}^L \equiv r_{t+1}^L - r_t \approx \frac{D}{1 - \theta} y_t^L - \frac{D - 1}{1 - \theta} y_{t+1}^L - r_t \tag{2} \]

where

\[ D \equiv \frac{\partial P^L_t}{\partial y_t^L} = \frac{1}{1 - \theta} = \frac{C + 1}{C} \tag{3} \]

is the Macaulay duration when the perpetuity is trading at par and \( y_t^L \) is the yield-to-maturity at time \( t \).

The risky asset is available in exogenous but time-varying fixed supply \( q_t \), where

\[ q_t = q_0 + q_1 s_t \tag{4} \]

and where the dynamics of \( s_t \) follow

\[ s_{t+1} = \rho_s s_t + \varepsilon_{s,t+1} \tag{5} \]

where \( \text{Var}_t [\varepsilon_{s,t+1}] = \sigma^2_s \), and \( \text{Cov}(\varepsilon_{r,t+1}, \varepsilon_{s,t+1}) = 0 \).

There is a unit mass of investors, each with risk tolerance \( \tau \). The investors can earn a risky return, \( r_{t}^L \), from the risky perpetual bond, or invest risklessly at rate \( r \). We assume that investors have mean-variance preferences over one-period portfolio returns. Mean-variance optimization implies that the demand for the risky asset, \( b_t^* \), is given by

\[ b_t^* = \underset{b_t}{\arg \max} \left\{ b_t E_t \left[ r_{x,t+1}^L \right] - \frac{(b_t)^2}{2\tau} \text{Var}_t \left[ r_{x,t+1}^L \right] \right\} \tag{6} \]

Equation (6) says that investor demand for the risky asset increases with excess return and risk tolerance, and decreases with the variance of excess return. Investor demand for the perpetual bond can be written as

\[ b_t^* = \tau \frac{E_t \left[ \sum_{j=1}^k r_{x,t+1} \right]}{\text{Var}_t \left[ \sum_{j=1}^k r_{x,t+1} \right]} \tag{7} \]
From equations (2), (7), and the market clearing condition \( (b_t^* = q_t) \), we obtain the risk premium of the risky asset, \( E_t[r_{x_{t+1}}^L] \), which is

\[
E_t[r_{x_{t+1}}^L] = \tau^{-1} q_t \left( \frac{\theta}{1 - \theta} \right)^2 \text{Var}_t[y_{t+1}^L] \quad (8)
\]

Rewriting the yield of the risky asset from (2) and rearranging (7), we can see that the equilibrium yield on the perpetual bond is a weighted average of expected future short rates and risk premia

\[
y_{t+1}^L = (1 - \theta)^\infty \sum_{i=0}^\infty \theta^i E_t[r_{t+i}] + \tau^{-1} b_{t+1}^* V_{rx}^* \quad (9)
\]

where \( V_{rx}^* = \left( \frac{\theta}{1 - \theta} \right)^2 \text{Var}_t[y_{t+1}^L] \) is one-period return volatility in equilibrium from (8). Using the market clearing condition \( (b_{t+1}^* = q_0 + q_1 s_{t+1}) \) and AR(1) dynamics, we can decompose the yield into

\[
y_{t+1}^L = \tau + \left( \frac{1 - \theta}{1 - \rho_s \theta} \right) (r_t - \tau) + \tau^{-1} V_{rx}^* q_0 \quad (10)
\]

Similarly, from (8) and (3), we can decompose the risk premium into

\[
E_t[r_{x_{t+1}}^L] = \frac{\tau^{-1} q_0 (D - 1)^2}{\sigma_y^2} + \frac{\tau^{-1} q_1 s_t (D - 1)^2}{\sigma_y^2} \quad (11)
\]

Equation (10) shows that yields are more sensitive to movements in short rates when the short-rate process is more persistent \( (\rho_r \text{ is larger}) \) and yields are more sensitive to movements in bond supply when the supply process is more persistent \( (\rho_s \text{ is larger}) \). Furthermore, since \( \theta = 1/(1 + C) = 1 - D^{-1} \), each of these sensitivities is greater when the bonds Macaulay duration is higher (i.e., when the coupon \( C \) is smaller).

Equation (39) clarifies the relationship between risk premia and supply shocks. When a positive supply shock is delivered to the market, \( s_t \) increases and the conditional risk premia increase as well. The magnitude of the change in the risk premia is inversely proportional to the risk tolerance and proportional to the size of the shock, the square of the duration (minus 1) and yield volatility. If the shock is permanent \( (\rho_s \approx 1) \), equation (35) and (39) dictate that \( s_t \) and the risk premia be permanently affected. If the shock is temporary \( (\rho_s < 1) \), \( s_t \) will revert to zero and the risk premia will decline to the unconditional level over time.

### 2.2 Two Asset Model With Slow-Moving Investors

We now develop the full model with two markets that are partially segmented as a result of slow-moving effects of cross-market investors. This model enables us to explore price dynamics following unanticipated and anticipated shocks to the supply of the asset. The model also allows us to explore
the pricing of supply risk by risk averse arbitrageurs as in Delong et al (1990), Vayanos and Vila (2009), Garleanu, Pedersen, and Potoshman (2009), and Greenwood and Vayanos (2013).

The model rests on the central premise of limited arbitrage theory that the marginal investor in any asset is a specialist with a large undiversified exposure to that asset. However, following Duffie (2010), we assume that there are a group of slow-moving investors who attempt to integrate these segmented markets at longer horizons. As a result, markets will be more segmented in the short-run and more integrated in the long-run. However, to the extent that integrating these markets is risky and the risk-tolerance of generalists is limited, these segmented markets will not be perfectly integrated even in the long-run.

2.2.1 Markets

There are two long-term fixed income assets, A and B. The first asset, denoted A, is default–free and is only exposed to interest rate risk. Thus, A is the same as the risky asset in the benchmark model developed above. In market B, we introduce a new source of risk.

Like the bond in market A, the bond in market B pays a promised coupon payment of C each period. However, B is subject to cash-flow risk. Specifically, assume that the realized coupon payment is \( C_t = C^L_t / Z_t \) where \( Z_t > 0 \) is cash-flow risk realization. If \( Z_t \equiv 1 \), the perpetuity is default-free. By contrast, if \( Z_t \) is stochastic, the perpetuity is defaultable with high realizations of \( Z_t \) corresponding to larger default losses at time \( t \). The gross return on the long-term security is \( 1 + R^L_{t+1} = (P^L_{t+1} + C_{t+1}) / P^L_t \). With the Campbell-Shiller (1988) approximation of the log return, we can write the holding period log return on asset B as

\[
\begin{align*}
rt_{t+1}^B &= \frac{1}{1 - \theta_B} y^B_t - \frac{\theta_B}{1 - \theta_B} y^B_{t+1} - (1 - \theta_B) z_{t+1} \\
\end{align*}
\]

where \( \theta_B = 1 / (1 + C_B) \). By comparison, the return of market A asset is

\[
\begin{align*}
rt_{t+1}^A &= \frac{1}{1 - \theta_A} y^A_t - \frac{\theta_A}{1 - \theta_A} y^A_{t+1} \\
\end{align*}
\]

where \( \theta_A = 1 / (1 + C_A) \). The additional \( z_{t+1} \) term in equation (12) reflects the cash-flow risk that is specific to market B. The variance of \( z_{t+1} \) determines the degree of similarity between assets A and B.

The process for the short rate \( r_t \) is the same as in equation (32). We assume that the cash-flow process follows

\[
zt_{t+1} = \bar{z} + \rho z (z_t - \bar{z}) + \varepsilon_{z,t+1} \\
\]

where \( Var_t [\varepsilon_{z,t+1}] = \sigma^2_z \). A high value of \( \sigma^2_z \) indicates less substitutability between markets A and B.

The two markets are subject to different supply shocks. Specifically, the net supply that bond investors must hold in A is

\[
q_t^A = q_0^A + q_1^A s_t^A \\
\]
where
\[ s_t^A = \rho s_t^A + \epsilon_{s,t+1}^A \]  
(16)
Similarly, the net supply that bond investors must hold in B is
\[ q_t^B = q_0^B + q_1^B s_t^B \]  
(17)
where
\[ s_{t+1}^B = \rho s_t^B + \epsilon_{s,t+1}^B. \]  
(18)
For simplicity we assume that \( \epsilon_{r,t+1}, \epsilon_{z,t+1}, \epsilon_{s,t+1}^A, \) and \( \epsilon_{s,t+1}^B \) are mutually orthogonal.

2.2.2 Market Participants

There is a unit measure of investors, each with risk tolerance \( \tau \). There are three types of investors who are distinguished by their ability to transact in different markets and by the frequency with which they can rebalance their portfolios. Fast-moving specialists in A allocate between short-rate and long-term asset in market A each period. They are present in mass \( p_A \) and their demand for asset A is denoted as \( b_t^A \). Fast-moving specialists in B allocate between short-rate and long-term asset in the B market each period. They are present in mass \( p_B \) and their demand for asset B is denoted as \( b_t^B \). Finally, there is a group of slow-moving generalists who allocate between both markets every \( k \) periods. Slow-moving generalists are present in mass \( 1 - p_A - p_B \). By varying \( p_A \) and \( p_B \) we ask how price dynamics change as we vary the mix of specialists and generalists, holding constant total investor risk tolerance.

Fast-moving specialists in A and B have mean-variance preferences over 1-period portfolio returns and their demands are described by (6) and (7). Slow-moving generalist investors have mean-variance preferences over the \( k \)-period cumulative return since they can only adjust every \( k \) periods. We let \( d_t^A \) and \( d_t^B \) denote the demand of the generalists who are active at time \( t \) in market A and B. Generalist investors choose their risky asset holdings to solve

\[
\max_{d_t^A, d_t^B} \left\{ -\frac{1}{2\tau} \left( d_t^A E_t \left[ \sum_{i=1}^{k} r_{t+i}^A \right] + d_t^B E_t \left[ \sum_{i=1}^{k} r_{t+i}^B \right] + E_t \left[ \sum_{i=0}^{k-1} r_{t+i} \right] \right) + \frac{(d_t^A)^2}{2} \text{Var}_t \left[ \sum_{i=1}^{k} r_{t+i}^A \right] + \frac{(d_t^B)^2}{2} \text{Var}_t \left[ \sum_{i=1}^{k} r_{t+i}^B \right] + 2d_t^A d_t^B \text{Cov}_t \left[ \sum_{i=1}^{k} r_{t+i}^A, \sum_{i=1}^{k} r_{t+i}^B \right] + 2d_t^A \text{Cov}_t \left[ \sum_{i=1}^{k} r_{t+i}^A, \sum_{i=0}^{k} r_{t+i} \right] + 2d_t^B \text{Cov}_t \left[ \sum_{i=1}^{k} r_{t+i}^B, \sum_{i=0}^{k} r_{t+i} \right] \right\}.
\]  
(19)

Equation (19) says that generalist investors demand more risky assets when the risk premia are high and when risk tolerance is high. They demand less risky assets when the cumulative excess returns over the \( k \)-periods between re-allocations, \( \sum_{i=1}^{k} r_{t+i}^A \) and \( \sum_{i=1}^{k} r_{t+i}^B \), are volatile, comove positively with one another, and/or comove positively with risk-free rate.
2.2.3 Equilibrium Yields

In market $A$, there are mass $p_A$ of fast-moving specialists, each with demand $b_t^A$, and mass $(1 - p_A - p_B)k^{-1}$ of slow-moving generalists who are active in period $t$, each with demand $d_t^A$. These investors must accommodate the active supply, which is the total supply $q_t^A = q_0^A + q_1^A s_t^A$ less any supply held off the market by inactive generalist investors, $(1 - p_A - p_B)k^{-1} \sum_{j=1}^{k-1} d_{t-j}^A$.

Thus, market clearing takes the form

$$(1 - p_A - p_B)k^{-1}d_t^A + p_A b_t^A = q_t^A - (1 - p_A - p_B)k^{-1} \sum_{i=1}^{k-1} d_{t-i}^A$$

(20)

The market clearing condition for $B$ is analogous.

We conjecture an equilibrium involving the state vector $x_t$, which informs the deviations from steady state at time $t$ for short-rate, cash-flow realization, supply shocks and past slow-generalist holdings of market $A$ and $B$ assets.

We conjecture that long-term yields in market $A$ and $B$ are

$$y_t^A = \alpha_{A0} + \alpha'_{A1} x_t$$

(21)

$$y_t^B = \alpha_{B0} + \alpha'_{B1} x_t$$

(22)

and that the demands of slow-moving generalists are

$$d_t^A = \delta_{A0} + \delta'_{A1} x_t$$

(23)

$$d_t^B = \delta_{B0} + \delta'_{B1} x_t$$

(24)

These assumptions imply that the state vector follows an AR(1) process. Critically, the transition matrix $\Gamma$ is a function of generalist demand so we write $\Gamma = \Gamma(\delta_{1A}, \delta_{1B})$.

The yields take same basic form as (9) with only specialist investors. However, $V_{rx}^i$ and $E_t [b_{t+i}^*]$ differ from previously as we have slow-generalist investors (i.e. $1 - p_A - p_B > 0$). For market $A$, the yield is

$$y_t^A = \underbrace{\mathbb{E}[\text{Expected future short rates}]}_{\mathbb{E}[\text{Unconditional term premia}]} + \underbrace{\left(1 - \frac{\theta_A}{1 - \rho_r \theta_A}\right) (r_t - \tau)}_{\mathbb{E}[\text{Conditional term premia}]} + \underbrace{(p_A \tau)^{-1} V_{rx}^{A*} (q_{A0} - (1 - p_A - p_B) \delta_{A0})}_{\mathbb{E}[\text{Unconditional term premia}]} + \underbrace{(p_A \tau)^{-1} V_{rx}^{A*} (1 - \theta_A) \sum_{i=0}^{\infty} \theta_A^i E_t \left[ q_{A1} s_{t+i}^A - (1 - p_A - p_B) k^{-1} \sum_{j=0}^{k-1} \delta_{t+i-j}^A \right]}_{\mathbb{E}[\text{Conditional term premia}]$$

(25)
In contrast, the yield for market $B$ asset has three extra terms relating to cashflow risk

$$y_t^B = [\text{Expected future short rates}] + [\text{Expected future defaults}]$$
$$+ [\text{Unconditional term premia}]_B + [\text{Conditional term premia}]_B$$
$$+ [\text{Unconditional default premia}] + [\text{Conditional default premia}]$$ (26)

Note that the (un)conditional term premia in market $A$ does not equal to the (un)conditional term premia in market $B$ because integrating the two markets is risky. The generalists are faced with cash-flow shocks (when $\sigma_z^2 > 0$) and differentiated supply shocks (when $\sigma_{sA}^2, \sigma_{sB}^2 > 0$).

### 2.3 Degree of Market Integration

Consider any pair of markets. The degree of integration is driven by two parameters $(1 - p_A - p_B)$ and $k$. $(1 - p_A - p_B)$ is the fraction of agents who are generalists able to allocate capital between the two markets. This determines the long-run degree of integration. For instance, if $(1 - p_A - p_B) \approx 1$, markets will be perfectly integrated in the long-run even even if $k$ is very large. $k$ indexes the speed with which generalist capital can flows between these markets. Thus, $k$ determines the short-run degree of integration.

Naturally, perfect integration is $(1 - p_A - p_B) = k = 1$. In this case, the model collapses to a mult-asset analog of Vayanos and Vila (2009) and Greenwood and Vayanos (2014). Perfect segmentation is $(1 - p_A - p_B) = 0$ and/or $k = \infty$. In this case, the model collapses to two completely unrelated markets with no supply spillovers. This is the radical view of financial markets put forth in Krishnamurthy and Vissing-Jorgensen (2013).

Often times, when $(1 - p_A - p_B)$ is high, $k$ is low and vice versa: generalists have expertise in both markets and are able to rapidly move between them. For instance, US Treasury and US Agency bonds are often managed by the same individual or team within large institutions. As a result, we would expect Treasury supply shocks to be rapidly transmitted to Agency debt.

However, there are important cases in which both $(1 - p_A - p_B)$ and $k$ are high. Consider stocks and bonds. The two assets are held significantly by diversified generalists precisely because they offer good diversification in combination, which is likely to be the case when the assets are exposed to very different fundamental shocks. In the language of our model, stocks are a high $\sigma_z^2$ version of market $B$ if we take Treasuries as market $A$.

### 3 Dynamics of Securities Prices

We start by describing pricing dynamics. We then examine how these dynamics depend on parameters of the model. We use a common set of parameter values across these simulations, which we lists in Table 1. Whenever possible, we calibrate these parameters under the assumption that
the U.S. Treasury market is market $A$ and that the corporate bond market is market $B$. As a starting point, we choose $k = 4$. That is, one-fourth of the generalists re-allocate their portfolio immediately.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_A, p_B$</td>
<td>Fraction of market participants that are specialists in $A$ and $B$ respectively</td>
<td>0.45</td>
</tr>
<tr>
<td>$k$</td>
<td>Number of periods for generalists to completely reallocate portfolio; inverse of the fraction of generalists with immediate response to shocks</td>
<td>4</td>
</tr>
<tr>
<td>$\sigma_z^2$</td>
<td>Cash-flow risk in market $B$; calibrated as the variance of corporate default rate</td>
<td>9</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Mean short rate (in %)</td>
<td>4</td>
</tr>
<tr>
<td>$\sigma_r^2$</td>
<td>Variance of short rate</td>
<td>1.7</td>
</tr>
<tr>
<td>$\bar{z}$</td>
<td>Mean corporate default rate (in %)</td>
<td>3</td>
</tr>
<tr>
<td>$\rho_r, \rho_z$</td>
<td>Persistences of short-rate and cash-flow risk processes</td>
<td>0.85</td>
</tr>
<tr>
<td>$\sigma_{sA}^2, \sigma_{sB}^2$</td>
<td>Variance of supply shocks</td>
<td>0.4</td>
</tr>
<tr>
<td>$D_A, D_B$</td>
<td>Market duration (in years)</td>
<td>5</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Risk tolerance</td>
<td>30</td>
</tr>
<tr>
<td>$q_0A, q_0B, q_{1A}, q_{1B}$</td>
<td>Steady state supply levels and coefficients on $s^A_t, s^B_t$</td>
<td>1</td>
</tr>
<tr>
<td>$\rho_{sA}, \rho_{sB}$</td>
<td>Persistences of supply shocks</td>
<td>1 (0.999)</td>
</tr>
</tbody>
</table>

Table 1: Parameters for simulations. Parameters are calibrated when possible using annual data with the U.S. Treasury market as market $A$ and the U.S. corporate bond market as market $B$. Unless otherwise specified, the parameters above are used in the simulations for security price dynamics and comparative statics.

### 3.1 Unanticipated Permanent Supply Shocks

We start by considering the impact of a permanent, unanticipated supply shock. Figure 1 illustrates the yield, return, and demand dynamics of two fixed-income markets in presence of a single unexpected positive supply shock that doubles the original supply of $A$ in period 10. We parametrize the simulation with $k = 4$, which means that a quarter of generalists reallocate their portfolio instantaneously and full reallocation is completed over 4 periods. Returns and yields of both markets react instantaneously. The top-left panel shows that $E[r x_A]$ overreacts, reaching a peak of 2.1% before ultimately falling back to a long-run level of 1.8%. $E[r x_B]$ under-reacts initially and rises...
slowly to a long-run level of 1.9%.

As shown in the bottom left-panel, the amount of overreaction/under-reaction is less in yield space since the yield is a weighted average of future returns. Market A yield overreacts by 11% of the total long-run impact and market B yield under-reacts by 17% of the total long-run impact.

The top right panel shows the demand dynamics of market participants. With initial supply shock in market A, both specialist demand in A \( (b_A) \) and generalist demand in A \( (d_A) \) spike upwards. In order to facilitate the increase in A demand, generalists reduce their portfolio holdings in market B \( (d_B) \). This reduction in generalist B holdings is motivated by a need to reduce the common short-rate risk across both markets. Thus, the generalists’ reduction of \( d_B \) is a hedge against their increase in \( d_A \). Lastly, specialists in market B step up their demand \( (b_B) \) to fill in the void left by the generalists. As more and more generalists wake up to reallocate their portfolio in the subsequent periods, the instantaneous demands for A \( (bA \) and \( dA) \) decrease slowly toward the long-run levels. The overreaction in demand in this simulation is illustrative of the relative steepness of short-run demand curve and flatness of the long-run demand curve.

The bottom-right panel shows the active supply of A and B available to market participants each period exclusive of the amount that is inaccessible for immediate portfolio reallocation. Initial supply shock in A increases the active supply in A but has no effect in the active supply for B. This is because the slow-moving generalists have yet to reduce their holdings in market B. Over the following periods, generalists increase their holdings of A and reduce their holdings of B. Therefore, the active supply in A gradually decreases while the active supply in B gradually increases.

Figure 2 shows that the yield spread between the two markets compresses in the long-run given more supply of A. This compression of yield spread is driven by the convergence of the risk premia between the two markets. The amount of overreaction observed in the yield spread is even larger (35%) due to a combination of overreaction of the yield A and underreaction of the yield B.
Figure 1: Impact of unexpected supply shock in primary and secondary market. A shock of size 1 (effectively doubling the original supply) is delivered to market A in period 10.

Figure 2: Impact of unexpected supply shock on yield spread between A and B. Yield spread here is defined as the yield B minus the yield A.
3.2 Anticipated Permanent Supply Shocks

We next study the dynamics following the announcement at time $t = T$ of a future supply shock. The motivation behind this exercise stems from the fact that QE announcements often precede well before the actual implementation of the designated purchases. To mimic the announcement of a future supply shock, we assume that $q_t$ jumps up at $t = T$, but also increase the lagged demands of slow-moving investors so that the active float is unchanged. Specifically, letting $\varepsilon_t [x_t] = E_t [x_{t+j}] - E_{t-1} [x_{t+j}]$ for some $j \geq 0$ denote the innovation to some random process, the supply shock is such that

$$0 = \varepsilon_T [q_T] - (1 - p) k^{-1} \sum_{j=1}^{k-1} \varepsilon_T [d_{T-j}] = [q_1 e_s - (1 - p) k^{-1} 1_{(d)}] \varepsilon_T [x_T].$$

(27)

where $e_s$ is the basis vector with a 1 corresponding to $s_t$ and 0s elsewhere and $1_{(d)}$ is a vector with 1s corresponding to the elements indicating deviations of previous generalist demands from unconditional level $\delta_0$.

For simplicity, we assume there is no change in expectations of future short rates or defaults. That is, $\tau$ and $\zeta$ are constants.

Note that we can also allow for different announced paths of purchases. Specifically, by holding $\varepsilon_T [q_T]$ fixed, we can vary the path of announced supply shocks by varying $\{\varepsilon_T [d_{T-j}]\}_{j=1}^{k-1}$ so long as we hold $\sum_{j=1}^{k-1} \varepsilon_T [d_{T-j}] = (k/ (1 - p)) \varepsilon_T [d_{T-j}]$. By holding $\varepsilon_T [q_T]$ and $\sum_{j=1}^{k-1} \varepsilon_T [d_{T-j}]$ fixed, we hold fixed the cumulative supply shock, but simply alter the announced timing of the shock.

Exercises like this can be used to evaluate different strategies for communications such as might be made by the Fed or by a company seeking to repurchase stock. Specifically, holding fixed the size of the purchase, we can ask whether we get a larger bang for buck using a gradual purchase strategy, a rapid purchase strategy, etc.

Figure 3 characterizes the dynamics of an anticipated supply shock that is announced at period 5 for implementation at period 8. The top-left panel shows that risk premia do not change immediately after the announcement. The majority of changes in risk premia happen after the actual delivery of the shock. The bottom-left panel shows that yields increase on the initial announcement of the anticipated supply shock but they do not fully adjust to the long-run levels immediately. The top-right panel shows that the generalists demand $A$ ($d_A$) increases and generalist demand $B$ ($d_B$) decreases on the initial announcement. Because generalists can only adjust their portfolio slowly, they must adjust in the direction of the eventual supply shock. That is, the generalists must accommodate the eventual increase in supply in period 8 by slowly building up their holdings of asset $A$ starting in period 5. The gradual build up of generalist demand in $A$ is responsible for the slight decrease in $ErxA$ from period 5 to 7 before the upward jump in period 8 from the actual implementation. Similarly, the gradual reduction of generalist $B$ demand results in the slow rise of $ErxB$. In contrast to the generalists, the fast specialist demand in market $A$ ($b_A$) decreases initially then increases. This is because specialists can adjust quickly and thus they reduce their portfolio
Figure 3: Impact of anticipated supply shock in primary and secondary market. A 50% increase in the supply of asset A is announced in period 5 and delivered in period 8 holdings of A just before the positive supply shock and increase holding of A immediately after the supply shock. The specialists are essentially front-running the anticipated change in supply.

Figure 4 compares the magnitude and reaction time of anticipated and unanticipated shocks of the same magnitude in yield space. Both shocks achieve the same long-run impacts on yields and premia. In the case of an unanticipated shock, overreaction is larger in the primary market. In the case of an anticipated shock, overreaction is smaller, however the amount of time it takes to achieve the full impact is longer. This suggests a trade-off for a central bank between slower policy implementation and higher market volatility. Allowing for a long period of time between policy announcement and implementation results in less overreaction, less profits to the specialists, and therefore less overall cost of policy implementation. On the other hand, the price adjustment when the shock is pre-announced is more gradual. There would be a cost associated with the delayed delivery of designated policy in the real economy.
Figure 4: Comparison of anticipated and unanticipated shock impacts. A 50% increase in the supply of asset \( A \) is delivered. Solid lines show the yield dynamics when the supply shock is announced in period 5 and delivered in period 8. Dashed lines show the yield dynamics when the supply shock is delivered in period 5 without early announcement.

We can also use the model to describe more complex paths of policy announcements. Consider a closer representation of central bank asset purchase. An announcement is made regarding policy implementation over multiple future periods. This example is analogous to the Fed’s purchase of $40 billion of MBS each month in QE3. In Figure 5, asset purchase is announced in period 5 and is carried out from period 6 to 14. The top-left panel shows the paths of risk premia falling over gradually over the periods of policy implementation. The bottom-left panel shows the paths of yields also declining gradually. Substantial under-reactions in returns and yields are observed for the two markets at the time of the policy announcement in period 5. An event study during the time of the initial announcement would capture some market reaction but it would far understate the long-run impact. We rationalize the limited initial reaction by analyzing the paths of generalist and specialist demands in the top right panel. Generalists and specialists both reduce their demands for \( A \) in order to facilitate the overall reduction of supply. Generalists also increase the demand for \( B \) to partially replace their holdings of \( A \). However, this market integration occurs at a slow pace.
Figure 5: Multiple Anticipated Shocks. In periods 6 to 14, a supply reduction equivalent to 5% of the original supply is made each period in market A. The asset purchase schedule and amount is announced in period 5. $k$ is set to 10 to allow for anticipated shocks to be delivered up to 10 periods following announcement. $\tau$ is set to 40 for solution existence.
since only one-tenth of the generalists can reallocate between the two markets each period. Lastly, specialists in B increase their demand \((b_B)\) to fill in for the generalists that have left market B for market A.

### 3.3 Comparative Statics

We next perform a number of comparative statics on supply shocks. We vary the values of key parameters in the model to illustrate the effects of an unanticipated shock on the yields. For simplicity, we focus on comparative statics in the cases of unanticipated shock on the yields. For simplicity, we focus on comparative statics in the cases of unanticipated shocks, but most of the results and intuitions are similar in the cases of anticipated shocks.

#### 3.3.1 \(\tau\)

\(\tau\) is the overall risk tolerance (shared by both the specialists and the generalists). As we make market participants more risk tolerant (increase \(\tau\)), we are essentially re-scaling the model.

![Figure 6: Comparative Statics on \(\tau\).](image)

The top left and top right panels show the paths of conditional yields for market A and B respectively. The bottom two panels show the paths of conditional risk premia. A positive supply shock equivalent to 30% of the original supply is delivered to market A in period 2.

#### 3.3.2 \(p_A, p_B\)
$p_A, p_B$ indicate the relative fraction of specialists in market $A$ and $B$. In the case $p_A = p_B = 0.5$, the fraction of generalists, $1 - p_A - p_B$, is zero. The two markets are completely segmented and a supply shock in market $A$ is not transmitted to market $B$. In the case there are many generalists, the markets are well-integrated and we have over-reaction in both markets. This is similar to the results in Duffie (2011). The two markets behave essentially as one. In the case there are a small number of generalists, the yield of the asset in the primary market over-reacts to supply shock while the yield of the asset in the secondary market under-reacts.

**Figure 7: Comparative Statics on $p_A$, $p_B$.** A positive supply shock equivalent to 30% of the original supply is delivered to market $A$ in period 2. $\sigma_{sA}^2$ and $\sigma_{sB}^2$ are set to 0 to allow for the solutions to exist for the case $p_B = p_A = 0.5$.

### 3.3.3 Relative fraction of $p_A, p_B$

Holding the fraction of generalists the same ($p_A + p_B = 0.8$), as we increase the fraction of specialists in $A$ and decrease the fraction of specialist in $B$, we get less over-reaction in $A$. The $B$ market is mostly unaffected by the change in the relative fraction of market specialists.
Figure 8: Comparative statics on $p_A$ holding constant $p_A + p_B = 0.8$. A positive supply shock equivalent to 30% of the original supply is delivered to market $A$ in period 2. $\tau$ is set to 40 to allow for the solutions to exist.

3.3.4 $\sigma_z^2$

As we increase $\sigma_z^2$, we have a larger $Z$-specific risk factor. We have less long-run and short-run integration between the markets. If $B$ specific risk is large, then the market with the shock will have a larger peak (because there is less ability to integrate between the markets). However, there is not much effect on under-reaction in the market that does not receive the shock (ie, under-reaction is consistent across scenarios). In the limit when $\sigma_z^2 = 0$, the markets would be perfect substitutes and thus perfectly integrated, and long-term conditional risk premia would be identical in the two markets. What we are seeing when we reduce $\sigma_z^2$ is that the conditional risk premia in market $B$ are actually rising because they are making $B$ a better substitute for $A$ (note that in isolation, you would get the opposite effect!), thus equalizing risk premia across the two markets in the long run.

A large $\sigma_z^2$ is analogous to a risky bond with mostly cash-flow risk. With high $\sigma_z^2$, the increase in yield and risk premia on $B$ following a supply shock to $B$ is dramatic. We are increasing the net supply of the risk factor which is very risky. The small amount of $r$ risk shared between the two markets is not enough to induce substantial market integration.
Figure 9: Comparative Statics on $\sigma_z^2$ with supply shock in $B$. A positive supply shock equivalent to 30% of the original supply is delivered to market $B$ in period 2. $\tau$ is set to 45 to allow for the solutions to exist.

3.3.5 $k$

$k$ is the number of periods it takes for generalists to fully reallocate their portfolio. When $k = 2$, half the generalists reallocate their portfolio immediately, and the other half reallocates in the next period. As the speed of reallocation decreases ($k$ increases), over-reaction in the primary market and under-reaction in the secondary market become more pronounced. With $k = 10$, only one-tenth of the generalists can react immediately to supply shock. Short-run yield and expected return changes represent only a fraction of the long-run changes.
Figure 10: Comparative Statics on $k$. A positive supply shock equivalent to 30% of the original supply is delivered to market $A$ in period 2. $\tau$ is set to 40 to allow for the solutions to exist.

3.4 Within Market Integration

Thus far, we have only considered a single risky asset in each market. We can extend our analysis to allow for multiple assets in each market. In fixed income markets, it is natural to consider the multiple assets as bonds of different maturities in the same market. We have developed pricing kernels for market $A$ and market $B$ by considering the supply of a homogeneous perpetually outstanding security that proxies for market as whole. We can introduce a full zero-coupon curve for each market under the assumption that these zero-coupon bonds are in zero net supply. For instance, yield curves look like

$$
\begin{align*}
y_t^{A(n)} &= n^{-1} \sum_{i=0}^{n-1} E_t \left[ r_{t+i} + (1-\theta) (n-i) \cdot \tau^{-1} b_{t+i}^A V_A \right] \\
y_t^{B(n)} &= n^{-1} \sum_{i=0}^{n-1} E_t \left[ r_{t+i} + (1-\theta) z_{t+i+1} + (1-\theta) (n-i) \cdot \tau^{-1} b_{t+i}^B V_B \right]
\end{align*}
$$

where $V_A$ and $V_B$ are the variance of 1-period returns.

With this method of constructing yield curve, we analyze the impact of a supply shock on the term structure of interest rates. Figure 11 shows the yields of 2-, 5-, and 10-year zero-coupon bonds as well as the yields of the coupon-bearing perpetual bonds. Longer duration bonds show less under-/over- reaction dynamics as they are averages of expected returns for many future periods. In contrast, a two year bond reflects only the expected returns over the current and the next period,
Figure 11: Impact of supply shock on yield curve. A positive supply shock equivalent to 30% of the original supply is delivered to market A in period 10. The yields of 2-, 5-, and 10- year zero-coupon bonds as well as the yields of the coupon-bearing perpetual bonds are shown.
Figure 12: Spread between 10-year bond yield and 2-year bond yield. A positive supply shock equivalent to 30% of the original supply is delivered to market $A$ in period 10. The yield spreads between 10-year zero-coupon bonds and 2-year zero-coupon bonds are shown for markets $A$ and $B$. The under-/over- reaction in yields are more significant as a result. This suggests that event study is more appropriate for bonds with longer maturity.

The spread between the yields of a long maturity bond and a short maturity bond is informative of the overall change in the shape of yield curve. Figure 12 shows that the spreads between 10-year yields and 2-year yields widen (the yield curves steepen) for both markets as a positive supply shock is delivered to market $A$. This matches with the empirical observation of yield curve changes during announcements of tapering by the Fed as exemplified by the taper tantrum in the summer of 2013.

4 Applications

Our model allows us to study a range situations in which price spillover occurs due to shocks in supply. It is also informative of the statistical detectability of price spillovers.

Consider the effects of $\sigma_z$ on the detection of price spillover in QE event studies. When $\sigma_z$ is large relative to $\sigma_r$, our model suggests that the econometricians may have little power to detect this gradual shift in excess returns: there is simply too much cash-flow news noise in market $B$ to reliably detect cross-market spillovers using a handful of QE announcement events. The statistical power would increase with the number of events, but power is still decreasing in $\sigma_z$. More generally, our framework cautions against drawing negative conclusions about price spillovers based on a
handful of events for some pair of assets.

Figure 13 shows the detectability of yield reactions to an unanticipated supply shock in market A. In an environment with low short-rate volatility and low supply shock volatility, a supply shock can have statistically significant impact in the primary market but not in the secondary market. The parametrization of Figure 13 correspond to the low volatility environment during QE when the short-term rate is pinned down at zero and supply risk is low. During this period, risk-taking by speculators was low ($\tau$ is small) and mortgage convexity was not an important contributor of supply shocks (as all the pre-payment options were either already exercised or deep in-the-money). Supply shocks induced by the usual hedging demand from insurance companies and pension funds were also lower than normal due to low rate volatility. In such an environment, event studies would have detected significant impact in one market, e.g. the MBS market, and little spillover to the secondary market, e.g. the treasury market. Even though the impact in the secondary market exists, the confidence interval is much wider and thus it would not be detectable by conventional event studies.

![Figure 13: Event studies confidence interval.](image)

The yields of markets A and B and the 95% confidence intervals of the respective market yield are shown. A doubling of supply is implemented in market A in period 10. The following parameters are used: $\tau = 0.5, \sigma_{sA} = \sigma_{sB} = 0, \sigma_r = 0.01, \sigma_z = 1$

5 Conclusion

In real world financial markets, arbitrage is a highly specialized activity. Specialized arbitrageurs ensure that options on IBM stock are priced consistently. Different arbitrageurs ensure that two
similar maturity Treasury bonds trade at the same yield. Yet another set of traders ensure that
the price of the S&P 500 futures contract does not deviate significantly from the basket of its
underlying stocks. Even outside of pure arbitrage, the business of money management is highly
segmented across markets. The vast majority of funds in both Fidelity and Vanguard, for example,two of the world’s largest money managers, are in vehicles that are specialized in either debt or
equity markets, but not both.

While specialization brings many benefits, the boundaries of securities markets are tested when
markets receive large shocks. In this paper, we developed a model to describe securities prices when
shocks must draw in arbitrageurs from other markets. We use the model to study the process by
which capital flows go across markets, and how quickly and by what magnitude prices adjust in
different markets. We show that even when capital is perfectly mobile in the long run, securities
prices do not have to be integrated across markets. This is because market segmentation creates
its own set of risks.
6 References


Miscellaneous QE References (a few I copied from Mamaysky paper that may be relevant, we don’t need to reference all these)


A Imperfect Integration

Even in the long run, these markets will be imperfectly integrated. This is because generalists are risk averse. As a result, integrating partially segmented markets is risky for generalists, even at long horizons.

To see this, suppose that $k = 1$, so generalists are just as nimble as specialists. We have

$$
d_t^A = \left[ \frac{\tau E_t [r_{xt_{t+1}}^A]}{\sigma_t^2} \right],
$$

$$
d_t^B = \left[ \frac{\tau E_t [r_{xt_{t+1}}^B]}{\sigma_t^2} \right].
$$

Market clearing requires

$$
p_A \frac{\tau E_t [r_{xt_{t+1}}^A]}{\sigma_t^2} + (1 - p_A - p_B) \left( \frac{\tau E_t [r_{xt_{t+1}}^A]}{\sigma_t^2} - \frac{\tau E_t [r_{xt_{t+1}}^B]}{\sigma_t^2} \right) = q_A
$$

$$
p_B \frac{\tau E_t [r_{xt_{t+1}}^B]}{\sigma_t^2} + (1 - p_A - p_B) \left( \frac{\tau E_t [r_{xt_{t+1}}^B]}{\sigma_t^2} - \frac{\tau E_t [r_{xt_{t+1}}^A]}{\sigma_t^2} \right) = q_B
$$

or

$$
\begin{bmatrix}
E_t [r_{xt_{t+1}}^A] \\
E_t [r_{xt_{t+1}}^B]
\end{bmatrix}
= \frac{1}{\tau} \begin{bmatrix}
\frac{p_A}{\sigma_t^2} + (1 - p_A - p_B) \left( \frac{1}{\sigma_t^2} + \frac{1}{\sigma_z^2} \right) - \frac{1 - p_A - p_B}{\sigma_t^2 + \sigma_z^2} \\
- \frac{1 - p_A - p_B}{\sigma_t^2 + \sigma_z^2} \frac{p_B}{\sigma_t^2} + \frac{1 - p_A - p_B}{\sigma_z^2}
\end{bmatrix}^{-1} \begin{bmatrix}
q_A \\
q_B
\end{bmatrix}
$$

$$
= \frac{1}{\tau} \begin{bmatrix}
q_A (1 - p_A - p_B) \sigma_t^2 (\sigma_t^2 + \sigma_z^2) + q_B (1 - p_A - p_B) \sigma_t^2 (\sigma_z^2 + \sigma_t^2) + q_B p_B \sigma_t^2 \sigma_z^2 \\
(q_A + q_B) (1 - p_A - p_B) \sigma_t^2 (\sigma_t^2 + \sigma_z^2) + q_B (1 - p_A - p_B) (\sigma_t^2 + \sigma_z^2)^2 + q_B p_B \sigma_t^2 (\sigma_t^2 + \sigma_z^2)
\end{bmatrix}
$$

Note that as $p_A, p_B \to 0$, we have

$$
\begin{bmatrix}
E_t [r_{xt_{t+1}}^A] \\
E_t [r_{xt_{t+1}}^B]
\end{bmatrix}
= \begin{bmatrix}
\frac{\sigma_t^2}{\tau} (q_A + q_B) \\
\frac{\sigma_z^2}{\tau} (q_A + q_B) + \frac{\sigma_t^2}{\tau} q_B
\end{bmatrix}.
$$

These are the standard integrated-market equilibrium excess returns. Thus, in the general case, we have

$$
\begin{bmatrix}
E_t [r_{xt_{t+1}}^A] \\
E_t [r_{xt_{t+1}}^B]
\end{bmatrix}
= \begin{bmatrix}
\frac{\sigma_t^2}{\tau} (q_A + q_B) \\
\frac{\sigma_z^2}{\tau} (q_A + q_B) + \frac{\sigma_t^2}{\tau} q_B
\end{bmatrix}
+ \frac{1}{\tau} \frac{1}{(\sigma_t^2 + \sigma_z^2) (1 - p_A - p_B) + \sigma_z^2 p B} \begin{bmatrix}
p_B \sigma_t^2 \sigma_z^2 (q_A - p_A (q_A + q_B)) \\
\sigma_z^2 p A ((\sigma_t^2 + \sigma_z^2) q_B - p_B (\sigma_t^2 q_A + (\sigma_t^2 + \sigma_z^2) q_B))
\end{bmatrix}.
$$
And
\[
E_t \left[ r_{x_i}^B \right] - E_t \left[ r_{x_i}^A \right] = \frac{\sigma_z^2}{\tau} \left( q_B + p_{APB} \frac{q_B (\sigma_t^2 + \sigma_z^2)}{p_B} - \frac{q_A \sigma_z^2}{p_A} - \sigma_z^2 q_B \right) \frac{1}{(\sigma_t^2 + \sigma_z^2)(1 - p_A - p_B) + \sigma_z^2 p_{APB}}
\]

The key insight here is that, generically, the second term is zero so long-run equilibrium pricing differs from integrated market pricing. For instance, \( E_t \left[ r_{x_i}^A \right] = \frac{\sigma_t^2}{\tau} (q_A + q_B) \) if: (1a) \( \sigma_z^2 = 0 \) (so integration is riskless); (1b) \( p_B = 0 \) (no B specialists); or (1c) \( q_A (1 - p_A) = p_A q_B \). Similarly, \( E_t \left[ r_{x_i}^B \right] = \frac{\sigma_t^2}{\tau} (q_A + q_B) + \frac{\sigma_z^2}{\tau} q_B \) if (2a) \( p_A = 0 \) (no A specialists) or (2b) \( (\sigma_t^2 + \sigma_z^2)(1 - p_B) q_B = \sigma_z^2 p_{BqA} \). Combining (1c) and (2b), we get integrated market pricing if
\[
p_A = \frac{q_A}{q_B + q_A} \text{ and } p_B = \frac{(\sigma_t^2 + \sigma_z^2) q_B}{\sigma_t^2 q_A + (\sigma_t^2 + \sigma_z^2) q_B}.
\]

However, this is not possible since we have
\[
\frac{q_A}{q_B + q_A} + \frac{(\sigma_t^2 + \sigma_z^2) q_B}{\sigma_t^2 q_A + (\sigma_t^2 + \sigma_z^2) q_B} > 1
\]
for \( \sigma_z^2, \sigma_t, q_A, q_B > 0 \). However, it is possible to get one (1c) or (2b) to hold. Just not both.

Effect of changes in \( q_A \) and \( q_B \):
\[
\left[ \frac{\partial E_t \left[ r_{x_i}^A \right]}{\partial q_A} \right] = \left[ \frac{\sigma_t^2}{\tau} + \frac{1}{\tau (\sigma_t^2 + \sigma_z^2)(1 - p_B)} \right] + \frac{\sigma_t^2}{p_{APB} \sigma_t^2 \sigma_z^2} - \sigma_z^2 q_B \right)
\]
\[
\left[ \frac{\partial E_t \left[ r_{x_i}^B \right]}{\partial q_B} \right] = \left[ \frac{\sigma_t^2}{\tau} + \frac{1}{\tau (\sigma_t^2 + \sigma_z^2)(1 - p_B)} \right] + \frac{\sigma_t^2}{p_{APB} \sigma_t^2 \sigma_z^2} - \sigma_z^2 q_B \right)
\]

Thus, we see that an expansion in \( q_A \) has a larger effect on \( E_t \left[ r_{x_i}^A \right] \) than on \( E_t \left[ r_{x_i}^B \right] \) (i.e., \( \partial E_t \left[ r_{x_i}^A \right]/\partial q_A > \tau^{-1} \sigma_t^2 > \partial E_t \left[ r_{x_i}^B \right]/\partial q_A \)). Similarly, we see that \( \partial E_t \left[ r_{x_i}^B \right]/\partial q_B > \tau^{-1} \sigma_t^2 + \tau^{-1} \sigma_z^2 \) and \( \partial E_t \left[ r_{x_i}^A \right]/\partial q_B < \tau^{-1} \sigma_z^2 \). Formally, this means that \( r \)-risk is not priced in the same way across the two markets. Formally, this arises because there is not a security that allows generalists to isolate the \( r \)-risk component of the \( B \) security.

## B Calibration with Annual Data

Annual calibration of the model is used for the numerical results.

**Calibration of Treasuries:** \( \sigma_z^2 = 0 \quad \text{For simplicity, we start with US Treasuries so} \quad \sigma_z^2 = z_t = 0 \). Note that we can also set \( q_0 = q_1 = 1 \) without loss of generality. Thus, we have
\[
y_t^L = \left[ \tau + \frac{1 - \theta}{1 - \rho_r \theta} (r_t - \tau) \right] + \tau^{-1} V^z \left[ 1 + \frac{1 - \theta}{1 - \rho_s \theta} s_t \right]
\]
and 
\[ E_t [rx_{t+1}^L] = \tau^{-1} V^* [1 + s_t] \]

and 
\[ V^* = Var_t [rx_{t+1}^L] = \left( \frac{\theta}{1 - \rho_r \theta} \right)^2 \sigma_r^2 + \left( \frac{\theta}{1 - \theta a_s^*} \right)^2 \sigma_s^2, \]

where 
\[ a_s^* = \frac{1 - \sqrt{1 - 4A_a C_a^s}}{2A_a} \]

and 
\[ 0 = \left[ \tau^{-1} \sigma_s^2 \theta \frac{\theta}{1 - \theta - \rho_s \theta} \right] (a_s)^2 - a_s + \left[ \tau^{-1} \left( \frac{\theta}{1 - \rho_r \theta} \right)^2 \sigma_r^2 \frac{1 - \theta}{1 - \rho_s \theta} \right]. \]

Note that the realized excess return is 
\[ rx_{t+1}^L = (E_t [rx_{t+1}^L]) + (rx_{t+1}^L - E_t [rx_{t+1}^L]) \]
\[ = \tau^{-1} V^* [1 + s_t] + \frac{\theta}{1 - \rho_r \theta} \varepsilon_{r,t+1} + \frac{\theta}{1 - \theta a_s^*} \varepsilon_{s,t+1}. \]

If we estimate 
\[ rx_{t+1}^L = a + b \cdot s_t + e_{t+1} \]
under the assumptions, we should obtain \( a = b = \tau^{-1} V^* \) with 
\[ R^2 = \frac{Var [\tau^{-1} V^* s_t]}{Var [rx_{t+1}^L]} = \frac{Var [\tau^{-1} V^* s_t]}{Var [\tau^{-1} V^* s_t]} \]
\[ = \frac{(\tau^{-1} V^*)^2 \sigma_s^2}{(1 - \rho_s^2) V^* (1 - \rho_r^2) V^* + \tau^2} \]
\[ = \frac{\sigma_s^2 (1 - \rho_s^2)}{(1 - \rho_r^2) V^* + \tau^2} \]

**Duration:** \( \theta = 5. \)

The mean modified duration of the Barclay’s Treasury index has been 5.22 years since 1989.

This means we want to use \( C = 0.25 \) implying a duration in years of \( D = (1 + C)/C = 5. \) This implies a price elasticity of 5 with respect to the annualized yield.

**Short rate:** \( \tau = 4; \rho_r = 0.85(0.8617244 \text{ in the data}); \sigma_r^2 = 1.7(1.7030339 \text{ in the data}) \)

**Cash-flow risk:** To calibrate \( \frac{\sigma_s^2}{1 - \rho_s^2} \) note that the \% 
\[ LOSS = DEF \times LGD = \frac{1}{Z} \]
or
\[ Z = \frac{1}{\text{DEF} \times \text{LGD}} \]
so we have
\[ z = -\ln(\text{DEF}) - \ln(\text{LGD}). \]

with corporate default data, we get \( \bar{\tau} = 3; \sigma_z^2 = 9; \rho_z = 0.85 \) (assumed to be similar to \( r \) risk)

**Other Parameters:** \( p_A = p_B = 0.45; \sigma_{sA}^2 = \sigma_{dB}^2 = 0.4; \tau = 30; k = 4 \)

### C Math Appendix

#### C.1 Single Asset Special Case

We consider the market for a single long-term fixed income asset which is exposed to interest rate risk. We also consider the possibility that this long-term asset is exposed to cash-flow risk—e.g., default risk or pre-payment risk—in addition to interest rate risk. For simplicity, we model this asset as perpetuity with a possibly stochastic coupon payment. As we will see, allowing for exposure to some other risk factor besides interest rate risk will useful once we examine the two asset version of the model.

There is a unit mass of investors, each with risk tolerance \( \tau \). There are two types of investors who are distinguished by the frequency with which they can rebalance their portfolios. Fast-moving investors allocate between the short-rate and the long-term asset in each period. By contrast, slow-moving investors allocate between the long-term asset and the short rate every \( k \) periods. Fast-moving investors have mass \( p \) and we let \( b_t \) denote the demand of fast-moving investors. Slow-moving investors have mass \( 1 - p \) and we let \( d_t \) denote the demand of the slow-moving investors who are active at time \( t \). Thus, by varying \( p \) we ask how price dynamics change as we vary the mix of fast and slow-moving investors, *holding constant total investor risk tolerance*.

#### C.1.1 Approximating the Log Return on a Risky Perpetuity

Consider a perpetuity with a *promised* coupon payment of \( C \) each period. If the yield-to-maturity on this security is \( Y_t^L \), its price is given by
\[ P_t^L = C \times \sum_{j=1}^{\infty} (1 + Y_t^L)^{-j} = C/Y_t^L. \]
We assume that the realized coupon payment is \( C_t = C^L/Z_t \) where \( Z_t > 0 \) is cash-flow risk realization. If \( Z_t \equiv 1 \), the perpetuity is default-free. By contrast, if \( Z_t \) is stochastic the perpetuity is *defaultable* with high realizations of \( Z_t \) corresponding to larger default losses at time \( t \).

The gross return on the long-term security is
\[ 1 + r_{t+1}^L = (P_{t+1}^L + C_{t+1}) / P_t^L, \]
where \( P_{t+1}^L \) is the price tomorrow and \( C_{t+1} \) is the uncertain cash-flow generated tomorrow. Using the the Campbell-Shiller (1988) approximation of the log return, we obtain
\[ r_{t+1}^L \approx \kappa + \theta p_{t+1}^L + (1 - \theta) (c - z_{t+1}) - p_t^L \]
(28)
where $\theta \equiv 1/(1 + \exp(c - z - p)) < 1$ and $\kappa = -\log(\theta) - (1 - \theta) \log(\theta^{-1} - 1)$. Iterating forward, this implies that

$$p^L_t = \frac{\kappa}{1 - \theta} + c - \sum_{i=0}^{\infty} \theta^i E_t \left[ r_{t+i+1}^L + (1 - \theta) z_{t+i+1} \right]$$

(29)

Applying this approximation to promised cashflows and the promised yield-to-maturity, we have

$$p^L_t = \frac{\kappa}{1 - \theta} + c - \frac{1}{1 - \theta} y^L_t$$

(30)

Assuming that the steady state price is par ($P^L = 1$), we have $\overline{p} = 0$ and $\theta = 1/(1 + C/Z)$. Thus, bond duration is

$$D \equiv \frac{\partial p^L_t}{\partial y^L_t} = \frac{1}{1 - \theta} = \frac{C/Z + 1}{C/Z}$$

Since Macaulay duration is

$$-\frac{\partial p^L_t}{\partial y^L_t} = \frac{\partial P^L_t}{\partial Y^L_t} \frac{1 + Y^L_t}{P^L_t} = \frac{Y^L_t + 1}{Y^L_t}$$

this corresponds to the Macaulay duration when the perpetuity is trading at par ($Y^L_t = C/Z$).

Substituting (30) into (28), we obtain

$$r^L_{t+1} = \frac{1}{1 - \theta} y^L_t - \frac{\theta}{1 - \theta} y^L_{t+1} - (1 - \theta) z_{t+1} = y^L_t - (D - 1) (y^L_{t+1} - y^L_t) - D^{-1} z_{t+1}$$

(31)

Naturally, realized returns from $t$ to $t+1$ are low (i) when the initial yield at $t$ is low, (ii) when yields rise between $t$ and $t+1$ leading to a capital loss, or (iii) when the default realization is high at $t + 1$.\footnote{Check: Steady-state log yield is

$$\overline{y} = \kappa + (1 - \theta) c$$

$$= -\log(\theta) - (1 - \theta) \log \left( \frac{1 - \theta}{\theta} \right) + (1 - \theta) c$$

$$= - (1 - \theta) \log(1 - \theta) - \theta \log(\theta) + (1 - \theta) c$$

$$= - \left( 1 - \frac{Z}{Z+C} \right) \log \left( 1 - \frac{Z}{Z+C} \right) - \frac{Z}{Z+C} \log \left( \frac{Z}{Z+C} \right) + \left( 1 - \frac{Z}{Z+C} \right) \log(C)$$

$$= \log(Z + C) - \frac{Z}{Z+C} \log(Z)$$

$$\approx \log \left( \frac{C + Z}{Z} \right)$$

}  

C.1.2 Simple Model with Only Fast-Moving Investors

We first explore a simple version of the model that can be solved in closed form. Specifically, we solve the model for a single long-term asset in the case where all investors are fast-moving. This obtains as a limiting case of the more general model when $p = 1$. In this special case, our model is the discrete time analog of Vayanos and Vila (2009) and Greenwood and Vila (2014).
Let \( r_{L_{t+1}} = r^{L}_{t+1} - r_t \) denote the 1-period excess return of this long-term bond over the short-rate. We assume that fast-moving investors have mean-variance preferences over their 1-period portfolio returns. This implies that investor demand is

\[
b_t^* = \arg \max_{b_t} \left\{ b_t E_t \left[ r_{L_{t+1}} \right] - \frac{(b_t)^2}{2\tau} Var \left[ r_{L_{t+1}} \right] \right\}.
\]

Thus, we have

\[
b_t^* = \tau \frac{E_t \left[ r_{L_{t+1}} \right]}{Var \left[ r_{L_{t+1}} \right]} = \tau \frac{1}{\theta} \frac{y_t^L - E_t \left[ y_{t+1}^L \right]}{\theta} Var \left[ y_{t+1}^L \right] + (1-\theta) E_t \left[ z_{t+1} \right] - r_t
\]

Suppose that the long-term asset is available in exogenous fixed supply \( q_t \). Market clearing \( (b_t^* = q_t) \) then implies

\[
E_t \left[ r_{L_{t+1}} \right] = \tau^{-1} q_t \left( \frac{\theta}{1-\theta} \right)^2 Var \left[ y_{t+1}^L \right] + (1-\theta)^2 Var \left[ z_{t+1} \right] + 2\theta Cov \left[ y_{t+1}^L, z_{t+1} \right]
\]

or

\[
y_t^L = \theta E_t \left[ y_{t+1}^L \right] + (1-\theta) \left( r_t + (1-\theta) E_t \left[ z_{t+1} \right] + E_t \left[ r_{L_{t+1}} \right] \right) + \tau^{-1} q_t \left( \frac{\theta}{1-\theta} \right)^2 Var \left[ y_{t+1}^L \right] + (1-\theta)^2 Var \left[ z_{t+1} \right] + 2\theta Cov \left[ y_{t+1}^L, z_{t+1} \right]
\]

so the equilibrium yield is a weighted average of tomorrow’s yield and \( r_t + (1-\theta) E_t \left[ z_{t+1} \right] + E_t \left[ r_{L_{t+1}} \right] \). Now consider a stationary equilibrium where \( Var_t \left[ y_{t+1}^L \right] = \sigma_y^2 \), \( Var_t \left[ z_{t+1} \right] = \sigma_z^2 \), and \( Cov \left[ y_{t+1}^L, z_{t+1} \right] = \sigma_{yz} \). Iterating forward implies

\[
y_t = (1-\theta) \sum_{i=0}^{\infty} (\theta)^i E_t \left[ r_{t+i} + (1-\theta) z_{t+i+1} + \tau^{-1} q_{t+i} \left( \frac{\theta}{1-\theta} \right)^2 \sigma_y^2 + (1-\theta)^2 \sigma_z^2 + 2\theta \sigma_{yz} \right]
\]

To close the model, we assume that the short-rate follows an exogenous stochastic process

\[
r_{t+1} = \bar{r} + \rho_r (r_t - \bar{r}) + \varepsilon_{r,t+1} \tag{32}
\]

where \( Var_t \left[ \varepsilon_{r,t+1} \right] = \sigma_r^2 \). We also assume that the cash flow follows

\[
z_{t+1} = \bar{z} + \rho_z (z_t - \bar{z}) + \varepsilon_{z,t+1} \tag{33}
\]

where \( Var_t \left[ \varepsilon_{z,t+1} \right] = \sigma_z^2 \). Finally, we assume that the net supply of the long-term bond that investors must hold is

\[
q_t = q_0 + q_1 s_t \tag{34}
\]

We assume that the dynamics of \( s_t \) follow

\[
s_{t+1} = \rho_s s_t + \varepsilon_{s,t+1} \tag{35}
\]

36
where $\text{Var}_t[\varepsilon_{s,t+1}] = \sigma_{z}^2$. For simplicity, we will assume that $\varepsilon_{r,t+1}$, $\varepsilon_{z,t+1}$, and $\varepsilon_{s,t+1}$ are mutually orthogonal. However, it is straightforward to relax this assumption.

Taken together these assumptions imply

$$
y_{t}^{L} = (1 - \theta) \sum_{i=0}^{\infty} (\theta)^i \left[ \tau + (\rho_r)^i (r_t - \tau) + (1 - \theta) \left( \frac{\theta}{1 - \sigma} \right)^2 \sigma_{y}^2 + (1 - \theta)^2 \sigma_{z}^2 + 2\theta \sigma_{yz} \right]$$

$$= \left[ \tau + \frac{1 - \theta}{1 - \rho_r \theta} (r_t - \tau) \right] + (1 - \theta) \left[ \frac{\theta}{1 - \rho_z \theta} \rho_z (z_t - \tau) \right]$$

$$+ \left[ \frac{1 - \theta}{1 - \rho_z \theta} \rho_z \sigma_{yz} \right]$$

To solve the model, we simply need to determine the equilibrium values of $\sigma_{yz}$ and $\sigma_{y}^2$. Since $\varepsilon_{r,t+1}$, $\varepsilon_{z,t+1}$, and $\varepsilon_{s,t+1}$ are mutually orthogonal, (36) implies that

$$\sigma_{yz} = \text{Cov}_t [y_{t+1}^L, z_{t+1}] = \frac{(1 - \theta)^2}{1 - \rho_z \theta} \rho_z \sigma_{z}^2$$

Using this result, we see that the equilibrium value of $\sigma_{y}^2$, denoted $\sigma_{y}^{2*}$, must satisfy the following fixed point condition

$$\sigma_{y}^{2*} = \frac{(1 - \theta)^2}{(1 - \rho_r \theta)^2} \sigma_{r}^2 + \frac{\rho_z^2 (1 - \theta)^4}{(1 - \rho_z \theta)^2} \sigma_{z}^2 + \frac{q_1 \sigma_s}{\tau} \left( \frac{\theta}{1 - \rho_z \theta} \right)^2 \left( \frac{\theta}{1 - \rho_z \theta} \right)^2 \sigma_{y}^2 + \frac{(1 - \theta)^3}{1 - \rho_z \theta} \sigma_{r}^2 + \frac{(1 - \theta)^4}{1 - \rho_z \theta} \sigma_{z}^2$$

Equation (37) implied a quadratic equation in $\sigma_{y}^2$. Specifically, we want to find the roots $\sigma_{y}^2$ that satisfy

$$0 = \frac{A_{\sigma}}{\sigma_{r}^2} + \frac{B_{\sigma}}{\sigma_{z}^2} + \frac{C_{\sigma}}{\sigma_{y}^2}$$

$$= \frac{\left( \frac{q_1 \sigma_s}{\tau} \right)^2 \left( \frac{\theta}{1 - \rho_z \theta} \right)^2 \left( \frac{\theta}{1 - \rho_z \theta} \right)^2 \sigma_{y}^2 + \frac{(1 - \theta)^2}{(1 - \rho_z \theta)^2} \sigma_{r}^2 + \frac{(1 - \theta)^4}{(1 - \rho_z \theta)^2} \sigma_{z}^2}{\left( \frac{\theta}{1 - \rho_z \theta} \right)^2 \left( \frac{\theta}{1 - \rho_z \theta} \right)^2}$$

This has solutions

$$\sigma_{y}^{2*} = \frac{-B_{\sigma} \pm \sqrt{B_{\sigma}^2 - 4A_{\sigma}C_{\sigma}}}{2A_{\sigma}}$$

So a solution exists if $B_{\sigma}^2 - 4A_{\sigma}C_{\sigma} > 0$. A bit of tedious algebra shows that this is the case if

$$\tau > q_1 \sigma_s \sqrt{\left( \frac{\theta}{1 - \rho_z \theta} \right)^2 \sigma_{r}^2 + \left( \frac{\theta}{1 - \rho_z \theta} \right)^2 \sigma_{z}^2}$$

(38)
Equation (38) shows that when $\sigma_s > 0$, a solution is only guaranteed to exist for $\tau$ sufficiently large. Furthermore, we have two solutions condition (38) holds, one solution of this equation holds with equality, and no solutions if the reverse inequality holds. This is precisely as in Vayanos and Vila (2009) and Vayanos and Greenwood (2014). The reason no equilibrium exits if $\tau$ is too small is that, in the presence of supply shocks and limited risk bearing capacity, bonds become extremely risky for arbitrageurs and it is impossible to clear the market. However, if $\tau$ is large enough so that an equilibrium exists, there are two equilibria: one in which yields are highly sensitive to supply shocks and one in which yields are less sensitive to supply shocks. If yields are highly sensitive to shocks to the supply risk factor $s_t$, then bonds become highly risky for arbitrageurs. Hence, arbitrageurs absorb supply shocks only if they are compensated by large changes in yields, making the high sensitivity of yields to shocks self-fulfilling.\(^2\)

We can compute comparatives for $\sigma^2_y$. Specifically, comparative statics for a change in any exogenous parameter $\gamma$ follow from

$$\frac{\partial \sigma^2_y}{\partial \gamma} = -\frac{\frac{\partial A_\sigma}{\partial \gamma} (\sigma^2_y)^2 + \frac{\partial B_\sigma}{\partial \gamma} \sigma^2_y + \frac{\partial C_\sigma}{\partial \gamma}}{2A_\sigma \sigma^2_y + B_\sigma}$$

If we focus on the smaller root, we have $2A_\sigma \sigma^2_y + B_\sigma < 0$ so

$$\frac{\partial X^*}{\partial \gamma} \propto \frac{\partial A_\sigma}{\partial \gamma} (\sigma^2_y)^2 + \frac{\partial B_\sigma}{\partial \gamma} \sigma^2_y + \frac{\partial C_\sigma}{\partial \gamma}$$

For instance, one can show that

$$\frac{\partial \sigma^2_y}{\partial \left(\frac{q_1 \sigma_t}{T}\right)}^2 \propto \sigma^2_y - \frac{\left(\frac{(1-\theta)^2}{(1-\rho_\theta)}\sigma^2_r + \frac{(1-\theta)^2}{(1-\rho_z \theta)} \rho_z^2 \sigma^2_z\right)}{\left(\frac{q_1 \sigma_t}{T}\right)^2} > 0,$$

where the term in square brackets is the variance of $y$ when $\sigma^2_y = 0$.

Combining all of this we have

$$E_t [r_{x_{t+1}^L}] = \tau^{-1} q_0 \left( \left( \frac{\theta}{1-\theta} \right)^2 \sigma^2_y + (1-\theta)^2 \frac{1 + \theta \rho_z}{1-\rho_z \theta} \sigma^2_z \right)$$

Combining all of this we have

\[ E_t [r_{x_{t+1}^L}] = \tau^{-1} q_0 \left( \left( \frac{\theta}{1-\theta} \right)^2 \sigma^2_y + (1-\theta)^2 \frac{1 + \theta \rho_z}{1-\rho_z \theta} \sigma^2_z \right) \]

\[ + \tau^{-1} q_1 s_t \left( \left( \frac{\theta}{1-\theta} \right)^2 \sigma^2_y + (1-\theta)^2 \frac{1 + \theta \rho_z}{1-\rho_z \theta} \sigma^2_z \right) \]

\[ E_t [r_{x_{t+1}^L}] = \tau^{-1} q_0 \left( \left( \frac{\theta}{1-\theta} \right)^2 \sigma^2_y + (1-\theta)^2 \frac{1 + \theta \rho_z}{1-\rho_z \theta} \sigma^2_z \right) \]

\[ + \tau^{-1} q_1 s_t \left( \left( \frac{\theta}{1-\theta} \right)^2 \sigma^2_y + (1-\theta)^2 \frac{1 + \theta \rho_z}{1-\rho_z \theta} \sigma^2_z \right) \]
and

\[
y_t^L = \tau + (1 - \theta) \left( r_t - \tau \right) + (1 - \theta) \left( z_t - \tau \right) + \frac{1}{1 - \rho_z \theta} \rho_z (z_t - \tau)
\]

(40)

Expected future short rates

Expected future cash flow losses

Unconditional risk premia

Expected future conditional risk premia

In the case, where \( \sigma_z^2 = 0 \), equation (39) implies that

\[
E_t [r_{x_{t+1}^L}] = \tau^{-1} (q_0 + q_1 s_t) (D - 1)^2 \sigma_y^{2*}
\]

so the term premium depends inversely on risk tolerance and is proportional to the quantity of long-term bonds, the square of duration (minus 1), and yield volatility. When \( \sigma_z^2 > 0 \), the bond risk premium contains an additional term that captures the compensation investors demand for bearing cashflow risk in addition to discount rate risk. Furthermore, in the natural case we have \( \frac{\partial \sigma_y^{2*}}{\partial q_1} > 0 \), \( \frac{\partial \sigma_y^{2*}}{\partial \sigma_s^2} > 0 \), and \( \frac{\partial \sigma_y^{2*}}{\partial \tau} < 0 \).

Turning to yields, equation (40) shows that:

- Yields are more sensitive to movements in short rates when the short-rate process is more persistent (\( \rho_r \) is larger).
- Yields are more sensitive to movements in default rates when the default process is more persistent (\( \rho_z \) is larger).
- Yields are more sensitive to movements in bond supply when the supply process is more persistent (\( \rho_s \) is larger).

Furthermore, since \( \theta = 1/(1 + C) = 1 - D^{-1} \), each of these sensitivities is greater when the bonds Macaulay duration is higher (i.e., when the promised coupon \( C \) is smaller).

Alternate solution approach

Another solution approach is to note that we must have

\[
a_s = \tau^{-1} q_1 \left( \frac{\theta}{1 - \theta} a_r \right)^2 \sigma_r^2 + \left( \frac{\theta}{1 - \theta} a_s \right)^2 \sigma_s^2 + \left( \frac{\theta}{1 - \theta} a_z + (1 - \theta) \right)^2 \sigma_z^2 \cdot \frac{1 - \theta}{1 - \rho_s \theta}.
\]

Using \( a_r = \frac{1 - \theta}{1 - \rho_r \theta} \) and \( a_z = \frac{(1 - \theta)^2}{1 - \rho_z \theta} \rho_z \), we require

\[
0 = \left[ \frac{\tau^{-1} \sigma_s q_1}{1 - \theta (1 - \rho_s \theta)} \right] (a_s)^2 - a_s + \left[ \frac{\tau^{-1} q_1}{(1 - \rho_r \theta)^2} \sigma_r^2 + \left( \frac{1 - \theta}{1 - \rho_z \theta} \right)^2 \sigma_z^2 \right] \frac{1 - \theta}{1 - \rho_s \theta}.
\]

39
We have

$$a_s^* = \frac{1 \pm \sqrt{1 - 4A_{a_s}C_{a_s}}}{2A_{a_s}},$$

which collapses to

$$a_s = \tau^{-1}q_1 \left( \left( \frac{\theta}{1 - \rho_r \theta} \right)^{2} \sigma_r^2 + \left( \frac{\theta}{1 - \rho_s \theta} \right)^{2} \sigma_z^2 \right) \frac{1 - \theta}{1 - \rho_s \theta}$$

when $\sigma_s^2 = 0$.

In general, a solution exists when $1 - 4A_{a_s}C_{a_s} \geq 0$ or

$$1 > 4 \left( \frac{\sigma_s q_1}{\tau} \right)^2 \left( \left( \frac{\theta}{1 - \rho_s \theta} \right)^{2} \sigma_r^2 + \left( \frac{\theta}{1 - \rho_s \theta} \right)^{2} \sigma_z^2 \right)$$

or

$$\frac{\tau}{2} \geq \sigma_s q_1 \left( \left( \frac{\theta}{1 - \rho_s \theta} \right)^{2} \sigma_r^2 + \left( \frac{\theta}{1 - \rho_s \theta} \right)^{2} \sigma_z^2 \right)^{1/2}$$

We can compute comparatives for $a_s^*$. Specifically, comparative statics for a change in any exogenous parameter $\gamma$ follow from

$$\frac{\partial a_s^*}{\partial \gamma} = -\frac{\partial A_{a_s}}{\partial \gamma} (a_s^*)^2 + \frac{\partial C_{a_s}}{\partial \gamma}$$

If we focus on the smaller root, we have $2A_{a_s}a_s^* - 1 < 0$ so

$$\frac{\partial a_s^*}{\partial q_1} \propto a_s^*/q_1 > 0$$

For instance, one can show that

$$\frac{\partial a_s^*}{\partial \tau} \propto -a_s^*/\tau < 0, \frac{\partial a_s^*}{\partial \sigma_s^2} > 0,$$

and $\frac{\partial a_s^*}{\partial \rho_s} > 0$. These are the comparative statics one should expect: supply risk-premia become large when risk tolerance is low, supply is volatile, or supply is highly persistent.

**Summary**

$$a_r = \frac{1 - \theta}{1 - \rho_r \theta}$$

$$a_z = \frac{(1 - \theta)^2}{1 - \rho_z \theta}$$

and

$$a_s^* = \frac{1 - \sqrt{1 - 4A_{a_s}C_{a_s}}}{2A_{a_s}}$$

where
\[
0 = \left[ \tau^{-1} \sigma^2 q_1 \frac{\theta}{1 - \theta} \left( \frac{\theta}{1 - \rho_s \theta} \right) \right] (a_s)^2 - a_s + \left[ \tau^{-1} q_1 \left( \frac{\theta}{1 - \rho_r \theta} \right)^2 \frac{\theta}{1 - \rho_s \theta} \right] (a_s)^2 + \left( 1 - \frac{\theta}{1 - \rho_s \theta} \right) \left( 1 - \theta \right) \frac{\theta}{1 - \rho_s \theta} \].
\]

To make sense of the coefficient on \( z_t \), note that as \( \rho_z \to 1 \) a change in \( z_t \) is just like a change in the bond’s coupon each period and we have \( a_z = (1 - \theta) = 1/D \). If the bond is priced at par this permanent change in \( z_t \) has a one-for-one change in the bond’s price and a \( 1/D \) change in the bond’s yield.

Note that
\[
V^* = \text{Var}_t \left[ r_{x_{t+1}}^L \right]
= \left( \frac{\theta}{1 - \theta} \right)^2 \sigma^2_r + \left( \frac{\theta}{1 - \theta} a^*_s \right)^2 \sigma^2_s + \left( \frac{\theta}{1 - \theta} (1 - \theta) \right)^2 \sigma^2_s
= \left( \frac{\theta}{1 - \rho_s \theta} \right)^2 \sigma^2_r + \left( \frac{\theta}{1 - \rho_s \theta} a^*_s \right)^2 \sigma^2_s + \left( 1 - \theta \right) \left( \frac{1}{1 - \rho_s \theta} (1 + \theta (1 - \rho_z)) \right)^2 \sigma^2_s
\]

Thus, we have
\[
y_t^L = \left[ \bar{r} + \frac{1 - \theta}{1 - \rho_s \theta} (r_t - \bar{r}) \right] + (1 - \theta) \left[ \bar{z} + \frac{1 - \theta}{1 - \rho_s \theta} \rho_z (z_t - \bar{z}) \right]
+ \tau^{-1} q_0 V^*
+ \tau^{-1} q_1 V^* \left( 1 - \theta \right) \frac{1 - \rho_s \theta}{1 - \rho_s \theta} s_t
\]
and
\[
E_t \left[ r_{x_{t+1}}^L \right] = \tau^{-1} V^* (q_0 + q_1 s_t)
\]

Note that the realized excess return is
\[
r_{x_{t+1}}^L = (E_t \left[ r_{x_{t+1}}^L \right]) + (r_{x_{t+1}}^L - E_t \left[ r_{x_{t+1}}^L \right])
= \tau^{-1} V^* (q_0 + q_1 s_t) + \frac{\theta}{1 - \rho_s \theta} \varepsilon_{r,t+1} + \frac{\theta}{1 - \rho_s \theta} a^*_s \varepsilon_{s,t+1} + \frac{1 - \theta}{1 - \rho_s \theta} (1 + \theta (1 - \rho_z)) \varepsilon_{z,t+1}
\]
If we estimate
\[
r_{x_{t+1}}^L = a + b \cdot s_t + \epsilon_{t+1}
\]
under the assumptions, we should obtain \( a = \tau^{-1}V^*q_0 \) and \( b = \tau^{-1}V^*q_1 \) with

\[
R^2 = \frac{Var[E_t [rx_{t+1}^L]]}{Var[rx_{t+1}^L]} = \frac{Var[\tau^{-1}q_1V^*s_t]}{\tau} = \frac{(\tau^{-1}q_1V^*)^2 \sigma^2_{r^L}}{1-\rho^2_s} + V^*
\]

so that both \( b \) and \( R^2 \) goes to zero and \((\tau/q_1) \to \infty\).

### C.1.3 Single Asset with Both Fast- and Slow-Moving Investors

We now incorporate slow-moving capital effects into this model following Duffie (2010). This model enables us to explore price dynamics following unanticipated and anticipated shocks to the supply of the asset. The model also allows us to explore the pricing of supply risk by risk averse arbitrageurs as in Delong et al (1990), Vayanos and Vila (2009), Garleanu, Pedersen, and Poteshman (2009), and Greenwood and Vayanos (2013).

There is a unit mass of agents each with risk tolerance \( \tau \). There are two types of agents who are distinguished by the frequency with which the can rebalance their portfolios. Fast-moving investors allocate between short-rate and long-term asset in each period. By contrast, slow-moving investors allocate between the long-term asset and the short rate every \( k \) periods. Fast-moving investors have mass \( p \) and we let \( b_t \) denote the demand of fast-moving investors. Slow-moving investors have mass \( 1-p \) and we let \( d_t \) denote the demand of the slow-moving investors who are active at time \( t \). Thus, by varying \( p \) we ask how price dynamics change as we vary the mix of fast and slow-moving investors, holding constant total investor risk tolerance.

#### Fast-Moving Investors

As above, we assume that fast-moving investors investors have mean-variance preferences over their 1-period portfolio returns. This implies that investor demand is

\[
\hat{b}_t^* = \arg \max_{b_t} \left\{ b_t E_t [rx_{t+1}^L] - \frac{(b_t)^2}{2} Var[rx_{t+1}^L] \right\}.
\]

Thus, we have

\[
\hat{b}_t^* = \frac{E_t [rx_{t+1}^L]}{Var[rx_{t+1}^L]} = \frac{1}{1-\theta} y_t^L - \frac{\theta}{1-\theta} E_t [y_{t+1}^L] - (1-\theta) E_t [z_{t+1}] - r_t
\]

#### Slow-Moving Investors
Slow-moving investors allocate between the long-term asset and investing their money in a rolling investment at the short-rate every $k$ periods. We assume that these investors have mean-variance preferences over the cumulative return on their wealth over $k$ periods. If the investor invests $d_t$ units of their wealth in the long-term asset at $(1 - d_t)$ units in the short-term asset, their cumulative return is

$$d_t \cdot \sum_{j=1}^{k} r_{t+j} + (1 - d_t) \cdot \sum_{j=0}^{k-1} r_{t+j} = d_t \cdot \sum_{j=1}^{k} r_{x_{t+j}} + \sum_{j=0}^{k-1} r_{t+j}.$$ 

Thus, slow-moving investors choose their long-term bond holdings to solve

$$\max_{d_t} \left\{ E_t \left[ d_t \cdot \sum_{j=1}^{k} r_{x_{t+j}} + \sum_{j=0}^{k-1} r_{t+j} \right] - \frac{1}{2\tau} \text{Var}_t \left[ d_t \cdot \sum_{j=1}^{k} r_{x_{t+j}} + \sum_{j=0}^{k-1} r_{t+j} \right] \right\}$$

or

$$\max_{d_t} \left\{ d_t E_t \left[ \sum_{j=1}^{k} r_{x_{t+j}} \right] - \frac{(d_t)^2}{2\tau} \text{Var}_t \left[ \sum_{j=1}^{k} r_{x_{t+j}} \right] - \frac{d_t}{\tau} \text{Cov}_t \left[ \sum_{j=1}^{k} r_{x_{t+j}}, \sum_{j=0}^{k-1} r_{t+j} \right] \right\}$$

so we have

$$d_t^* = \tau \frac{E_t \left[ \sum_{j=1}^{k} r_{x_{t+j}} \right]}{\text{Var}_t \left[ \sum_{j=1}^{k} r_{x_{t+j}} \right]} - \frac{\text{Cov}_t \left[ \sum_{j=1}^{k} r_{x_{t+j}}, \sum_{j=0}^{k-1} r_{t+j} \right]}{\text{Var}_t \left[ \sum_{j=1}^{k} r_{x_{t+j}} \right]}$$

Note that slow-moving investors agents have a hedging motive: they dislike assets whose cumulative returns covary with the unknown cumulative return for investing in a series of 1-period bonds. (In practice, this plays only a minor role in our model.)

**Equilibrium Conjecture**

Suppose, for example, that $k = 5$. We look for a solution using the following state vector

$$x_t = \begin{bmatrix} r_t - \bar{r} \\ s_t \\ d_{t-1} - \delta_0 \\ d_{t-2} - \delta_0 \\ d_{t-3} - \delta_0 \\ d_{t-4} - \delta_0 \\ z_t - \bar{z} \end{bmatrix}.$$ \hspace{1cm} (41)

We conjecture that the equilibrium long-term yield is of the form

$$y_t^L = \alpha_0 + \alpha'_L x_t$$ \hspace{1cm} (42)

and that the demand of slow-moving investors is of the form

$$d_t = \delta_0 + \delta'_1 x_t$$ \hspace{1cm} (43)
Combined with our prior assumption on the dynamics of \( r_t, z_t, \) and \( s_t \), this implies that the state vector follows an AR(1) process. Critically, the transition matrix \( \Gamma \) is a function of inattentive demand so we write \( \Gamma = \Gamma (\delta_1) \). Specifically, we have

\[
x_{t+1} = \Gamma (\delta_1) x_t + \epsilon_{t+1}
\]

\[
= \begin{bmatrix}
\rho_r & 0 & 0 & 0 & 0 & 0 \\
0 & \rho_s & 0 & 0 & 0 & 0 \\
\delta_1 & \delta_2 & \delta_3 & \delta_4 & \delta_5 & \delta_7 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & \rho_z
\end{bmatrix}
\begin{bmatrix}
r_t - \bar{r} \\
s_t \\
d_{t-1} - \delta_0 \\
d_{t-2} - \delta_0 \\
d_{t-3} - \delta_0 \\
d_{t-4} - \delta_0 \\
z_t - \bar{z}
\end{bmatrix}
+ \begin{bmatrix}
\varepsilon_{r,t+1} \\
\varepsilon_{s,t+1} \\
0 \\
0 \\
0 \\
0 \\
\varepsilon_{z,t+1}
\end{bmatrix}
\]

where

\[
\text{Var}\left[\epsilon_{t+1}\right] = \Sigma = \begin{bmatrix}
\sigma_r^2 & 0 & 0 & 0 & 0 \\
0 & \sigma_e^2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \sigma_z^2 & 0 \\
0 & 0 & 0 & 0 & \sigma_z^2
\end{bmatrix}
\]

We adopt the convention that \( e_r \) is the basis vector with a 1 corresponding to \( r_t - \bar{r} \) and 0s elsewhere. Similarly, \( e_z \) is the basis vector with a 1 corresponding to \( z_t - \bar{z} \) and 0s elsewhere. And \( e_s \) is the basis vector with a 1 corresponding to \( s_t \) and 0s elsewhere. Finally, we denote \( \text{Cov}\left[ x_{t+k}, x_{t+j} | x_t \right] = \text{Cov}\left[ x_{t+k}, x_{t+j} | x_t \right] \) and note that

\[
\text{C}_{[t+i,t+j]} = \text{Cov}\left[ x_{t+k}, x_{t+j} | x_t \right] = \sum_{i=1}^{\min\{k,j\}} [\Gamma^{k-i}] \Sigma [\Gamma^{j-i}]'
\]

\subsection*{Model Solution}

\textbf{Fast-Moving Demand} \quad \text{Fast-moving demand is } b_t = \tau E_t \left[ r_{x_{t+1}^L} \right] / \text{Var}\left[ r_{x_{t+1}^L} \right]. \text{ Given our conjecture that }

\[
y_t^L = \alpha_0 + \alpha_1' x_t
\]

we have

\[
r_{x_{t+1}^L} = \frac{1}{1 - \theta} y_t^L - \frac{\theta}{1 - \theta} y_{t+1}^L - (1 - \theta) z_{t+1} - r_t
\]

\[
= (\alpha_0 - \bar{r} - (1 - \bar{z})) + \left( \frac{1}{1 - \theta} \alpha_1 - e_r \right)' x_t - \left( \frac{\theta}{1 - \theta} \alpha_1 + (1 - \theta) e_z \right)' x_{t+1}
\]
This implies that

$$E_t [r_{x_t^L}^L] = (\alpha_0 - \bar{r} - (1 - \theta) \bar{z}) + \left( \frac{1}{1 - \theta} \alpha_1 - e_r \right) x_t - \left( \frac{\theta}{1 - \theta} \alpha_1 + (1 - \theta) e_z \right)' \Gamma x_t$$

(48)

and

$$\text{Var}_t [r_{x_t^L}^L] = \left( \frac{\theta}{1 - \theta} \alpha_1 + (1 - \theta) e_z \right)' \Sigma \left( \frac{\theta}{1 - \theta} \alpha_1 + (1 - \theta) e_z \right)'$$

$$= \left( \frac{\theta}{1 - \theta} a_r \right)^2 \sigma_r^2 + \left( \frac{\theta}{1 - \theta} a_s \right)^2 \sigma_s^2 + \left( \frac{\theta}{1 - \theta} a_z + (1 - \theta) \right)^2 \sigma_z^2$$

(49)

Thus, fast-moving investor demand is given by

$$b_t = \left[ \frac{(\alpha_0 - \bar{r} - (1 - \theta) \bar{z})}{\left( \frac{\theta}{1 - \theta} a_r \right)^2 \sigma_r^2 + \left( \frac{\theta}{1 - \theta} a_s \right)^2 \sigma_s^2 + \left( \frac{\theta}{1 - \theta} a_z + (1 - \theta) \right)^2 \sigma_z^2} \right]$$

$$+ \left[ \frac{(\frac{\theta}{1 - \theta} \alpha_1 - e_r)' - \left( \frac{\theta}{1 - \theta} \alpha_1 + (1 - \theta) e_z \right)' \Gamma x_t}{\left( \frac{\theta}{1 - \theta} a_r \right)^2 \sigma_r^2 + \left( \frac{\theta}{1 - \theta} a_s \right)^2 \sigma_s^2 + \left( \frac{\theta}{1 - \theta} a_z + (1 - \theta) \right)^2 \sigma_z^2} \right] x_t$$

(50)

**Slow-Moving Demand**

We have

$$\sum_{j=1}^k r_{x_{t+j}^L} = \sum_{j=0}^{k-1} (y_{t+j}^L - r_{t+j} - (1 - \theta) z_{t+j+1}) - \frac{\theta}{1 - \theta} (y_{t+k} - y_{t})$$

$$= k (\alpha_0 - \bar{r} - (1 - \theta) \bar{z}) + \alpha_1 - e_r)' \left( \sum_{j=0}^{k-1} x_{t+j} \right)$$

$$- (1 - \theta) e_z' \left( \sum_{j=1}^k x_{t+j} \right) - \frac{\theta}{1 - \theta} (\alpha_1 x_{t+k} - \alpha_1' x_t)$$

(51)

This implies that

$$E_t \left[ \sum_{j=1}^k r_{x_{t+j}^L} \right] = k (\alpha_0 - \bar{r} - (1 - \theta) \bar{z}) + \alpha_1 - e_r)' (I - \Gamma)^{-1} (I - \Gamma^k) x_t$$

$$- (1 - \theta) e_z' (I - \Gamma)^{-1} (\Gamma - \Gamma^{k+1}) x_t - \frac{\theta}{1 - \theta} \alpha_1' (\Gamma^k - I) x_t$$

(52)

and

$$\text{Var}_t \left[ \sum_{j=1}^k r_{x_{t+j}^L} \right]$$

$$= \text{Var}_t \left[ (\alpha_1 - e_r - (1 - \theta) e_z)' (\sum_{j=1}^{k-1} x_{t+j}) - \left( \frac{\theta}{1 - \theta} \alpha_1 + (1 - \theta) e_z \right)' x_{t+k} \right]$$

$$= (\alpha_1 - e_r - (1 - \theta) e_z)' \left( \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} C_{[i,j]} \right) (\alpha_1 - e_r - (1 - \theta) e_z)$$

$$+ \left( \frac{\theta}{1 - \theta} \alpha_1 + (1 - \theta) e_z \right)' C_{[k,k]} \left( \frac{\theta}{1 - \theta} \alpha_1 + (1 - \theta) e_z \right)$$

$$- 2 (\alpha_1 - e_r - (1 - \theta) e_z)' \sum_{i=1}^{k-1} C_{[i,k]} \left( \frac{\theta}{1 - \theta} \alpha_1 + (1 - \theta) e_z \right)$$

(53)
and

\[
Cov_t \left[ \sum_{j=1}^{k} x_{t+j}, \sum_{j=0}^{k-1} r_{t+j} \right] = Cov_t \left[ (\alpha_1 - \mathbf{e}_r - (1 - \theta) \mathbf{e}_z)' \left( \sum_{j=1}^{k-1} x_{t+j} \right) - \left( \frac{\theta}{1 - \theta} \alpha_1 + (1 - \theta) \mathbf{e}_z \right)' \mathbf{x}_{t+k}, \mathbf{e}_r' \sum_{j=1}^{k-1} x_{t+j} \right] \tag{54}
\]

\[
= (\alpha_1 - \mathbf{e}_r - (1 - \theta) \mathbf{e}_z)' \left( \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} C_{[t+i,t+j]} \right) \mathbf{e}_r - \left( \frac{\theta}{1 - \theta} \alpha_1 + (1 - \theta) \mathbf{e}_z \right)' \sum_{i=1}^{k-1} C_{[t+i,t+k]} \mathbf{e}_r
\]

Combining these expressions, we see that the demand of fast-moving investors take the form

\[
d_t = \delta_0 + \delta_1^* x_t.
\]

**Market Clearing** There is a mass \( p \) of fast-moving investors each with demand \( b_t \). This is mass \((1 - p) k^{-1}\) of slow-moving investors who are active in period \( t \), each with demand \( d_t \). Thus, total active demand is \( pb_t + (1 - p) k^{-1} d_t \). The current supply that must be observed by these investors equals to total supply minus the purchases of slow-moving investors that are current hold off the market, \( q_t - (1 - p) k^{-1} \sum_{j=1}^{k-1} d_{t-j} \). Thus, to clear the market for long-term bonds in period \( t \), we require

\[
pb_t + (1 - p) k^{-1} d_t = q_t - (1 - p) k^{-1} \sum_{j=1}^{k-1} d_{t-j} \tag{55}
\]

Rewriting this as \( pb_t = q_t - (1 - p) k^{-1} \sum_{j=0}^{k-1} d_{t-j} \) we have

\[
\begin{align*}
\frac{pT}{\tau} \alpha_0 - \tau - (1 - \theta) \tau \\
\frac{2}{\tau p} \left( \frac{\theta}{1 - \theta} a_r \right)^2 \sigma_r^2 + \left( \frac{\theta}{1 - \theta} a_s \right)^2 \sigma_s^2 + \left( \frac{\theta}{1 - \theta} a_z + (1 - \theta) \right)^2 \sigma_z^2 \\
+ p \left( \frac{2}{\tau p} \left( \frac{1}{1 - \theta} \alpha_1 - \mathbf{e}_r \right)' - \left( \frac{\theta}{1 - \theta} \alpha_1 + (1 - \theta) \mathbf{e}_z \right)' \Gamma \right) \mathbf{x}_t \\
= (q_0 - (1 - p) \delta_0) + \left[ q_1 \mathbf{e}_s - (1 - p) k^{-1} (1_d + \delta_1) \right]' \mathbf{x}_t
\end{align*}
\]

Matching constants, we require

\[
\alpha_0 = \tau + (1 - \theta) \tau = \frac{\left( \frac{\theta}{1 - \theta} a_r \right)^2 \sigma_r^2 + \left( \frac{\theta}{1 - \theta} a_s \right)^2 \sigma_s^2 + \left( \frac{\theta}{1 - \theta} a_z + (1 - \theta) \right)^2 \sigma_z^2}{\tau p} (q_0 - (1 - p) \delta_0) \tag{56}
\]

Matching slope coefficients, we require

\[
\begin{align*}
\alpha_1 &= (1 - \theta) \left[ \mathbf{I} - \theta \Gamma \right]^{-1} \left[ \mathbf{e}_r + (1 - \theta) \Gamma' \mathbf{e}_z \right] \\
+ \frac{\left( \frac{\theta}{1 - \theta} a_r \right)^2 \sigma_r^2 + \left( \frac{\theta}{1 - \theta} a_s \right)^2 \sigma_s^2 + \left( \frac{\theta}{1 - \theta} a_z + (1 - \theta) \right)^2 \sigma_z^2}{\tau p_B} (1 - \theta) \left[ \mathbf{I} - \theta \Gamma \right]^{-1} \left[ q_1 \mathbf{e}_s - (1 - p) k^{-1} (1_d + \delta_1) \right]
\end{align*}
\]

**Solution details** Since \( \delta_{1r} = \delta_{1z} = 0 \), we will have

\[
a_r = (1 - \theta) \mathbf{e}_r' \left[ \mathbf{I} - \theta \Gamma \right]^{-1} \mathbf{e}_r = \frac{1 - \theta}{1 - \theta p_r}
\]
and

\[ a_z = (1 - \theta)^2 \varepsilon_k^2 \Gamma (I - \theta \Gamma)^{-1} \varepsilon_z = \frac{(1 - \theta)^2}{1 - \rho_z \bar{\theta}} p_z \]

We can probably expedite the Matlab solver by hardwiring these solutions.

To see that \( \delta_{1r} = \delta_{1z} = 0 \), let \( V^k = Var_t \left[ \sum_{j=1}^k r x_{t+1}^L \right] \) and \( V^1 = Var_t [r x_{t+1}^L] \), so that

\[ \alpha'_1 = (1 - \theta) [e'_r + (1 - \theta) e'_z \Gamma] [I - \theta \Gamma]^{-1} + \frac{V^1}{\tau p} (1 - \theta) \left[ q_1 \varepsilon_s - (1 - p) k^{-1} (1_{(d)} + \delta_1) \right]' [I - \theta \Gamma]^{-1} \]

and

\[ \frac{V^k}{\tau} \delta'_1 = (\alpha_1 - e_r)' (I - \Gamma)^{-1} (I - \Gamma^k) - (1 - \theta) e'_z (I - \Gamma)^{-1} (\Gamma - \Gamma^{k+1}) - \frac{\theta}{1 - \theta} \alpha'_1 (\Gamma^k - I) \]

Now since

\[ \left[ (I - \Gamma)^{-1} (I - \Gamma^k) - \frac{\theta}{1 - \theta} (\Gamma^k - I) \right] \]

\[ = \left[ (I - \Gamma)^{-1} + \frac{\theta}{1 - \theta} \right] (I - \Gamma^k) \]

\[ = \frac{1}{1 - \theta} [I (1 - \theta) + \theta (I - \Gamma)] (I - \Gamma)^{-1} (I - \Gamma^k) \]

\[ = \frac{1}{1 - \theta} [I - \theta \Gamma] (I - \Gamma)^{-1} (I - \Gamma^k) \]

we have

\[ \frac{V^k}{\tau} \delta'_1 \]

\[ = \frac{1}{(1 - \theta)} \alpha'_1 [I - \theta \Gamma] [I - \Gamma]^{-1} (I - \Gamma^k) - e'_r (I - \Gamma)^{-1} (I - \Gamma^k) - (1 - \theta) e'_z (I - \Gamma)^{-1} \Gamma (1 - \Gamma^k) \]

\[ = \frac{1}{(1 - \theta)} (1 - \theta) e'_r [I - \Gamma]^{-1} (I - \Gamma^k) - e'_r (I - \Gamma)^{-1} (I - \Gamma^k) - (1 - \theta) e'_z (I - \Gamma)^{-1} \Gamma (1 - \Gamma^k) \]

\[ + \frac{1}{(1 - \theta)} \frac{\sigma^2_p}{\tau p} (1 - \theta) \left[ q_1 \varepsilon_s - (1 - p) k^{-1} (1_{(d)} + \delta_1) \right]' [I - \theta \Gamma]^{-1} [I - \theta \Gamma] (I - \Gamma)^{-1} (I - \Gamma^k) \]

\[ = \frac{V^1}{\tau p} \left[ q_1 \varepsilon_s - (1 - p) k^{-1} (1_{(d)} + \delta_1) \right]' [I - \Gamma]^{-1} (I - \Gamma^k) \]

Letting \( V^k = Var_t \left[ \sum_{j=1}^k r x_{t+1}^L \right] \) and \( V^1 = Var_t [r x_{t+1}^L] \), we have

\[ p V^k \delta_1 = V^1 [I - (\Gamma^k)] [I - \Gamma']^{-1} \left[ q_1 \varepsilon_s - (1 - p) k^{-1} (1_{(d)} + \delta_1) \right] \]

or

\[ \delta_1 = V^1 \left[ p V^k + (1 - p) k^{-1} V^1 [I - (\Gamma^k)] [I - \Gamma']^{-1} \right] \left[ I - (\Gamma^k) \right] [I - \Gamma']^{-1} \left[ q_1 \varepsilon_s - (1 - p) k^{-1} 1_{(d)} \right] \]

From this expression, it is clear that \( \delta_1 \) must satisfy \( \delta_{1r} = \delta_{1z} = 0 \). In other words, in equilibrium shocks to \( r_t \) and \( z_t \) have no impact on either current fast-specialists or slow-generalist demand. Only shocks to \( s_t \) and changes in past slow-generalist demand impact current demands.

47
Equilibrium Yields and Excess Returns

Expected 1-period excess returns are

\[
E_t [r_{t+1}^L] = \tau^{-1} b_t^* \left( \frac{\theta}{1-\theta} a_r \right)^2 \sigma_r^2 + \left( \frac{\theta}{1-\theta} a_s \right)^2 \sigma_s^2 + \left( \frac{\theta}{1-\theta} a_z + (1-\theta) \right)^2 \sigma_z^2
\]

\[
= \frac{\left( \frac{\theta}{1-\theta} a_r \right)^2}{\tau_p} \sigma_r^2 + \frac{\left( \frac{\theta}{1-\theta} a_s \right)^2}{\tau_p} \sigma_s^2 + \left( \frac{\theta}{1-\theta} a_z + (1-\theta) \right)^2 \sigma_z^2 \left( q_0 - (1-p)\delta_0 \right)
\]

Equilibrium yields are

\[
y_t^L = (1-\theta) \sum_{i=0}^{\infty} (\theta)^i E_t \left[ r_{t+i} + (1-\theta) z_{t+i+1} + (p\tau)^{-1} b_{t+i}^* \left( \frac{\theta}{1-\theta} \sigma_y^2 + (1-\theta)^2 \sigma_z^2 + 2\theta \sigma_y^2 \right) \right]
\]

\[
= \tau + (1-\theta) e_r^\prime [I - \theta \Gamma]^{-1} x_t + (1-\theta) e_s^\prime [I - \theta \Gamma]^{-1} x_t
\]

\[
+ \frac{\left( \frac{\theta}{1-\theta} a_r \right)^2}{\tau_p} \sigma_r^2 + \frac{\left( \frac{\theta}{1-\theta} a_s \right)^2}{\tau_p} \sigma_s^2 + \left( \frac{\theta}{1-\theta} a_z + (1-\theta) \right)^2 \sigma_z^2 \left( q_0 - (1-p)\delta_0 \right)
\]

\[
+ \frac{\left( \frac{\theta}{1-\theta} a_r \right)^2}{\tau_p} \sigma_r^2 + \frac{\left( \frac{\theta}{1-\theta} a_s \right)^2}{\tau_p} \sigma_s^2 + \left( \frac{\theta}{1-\theta} a_z + (1-\theta) \right)^2 \sigma_z^2 \left[ q_1 e_s - (1-p)k^{-1} (1_{(d)} + \delta_1) \right] (1-\theta) [I - \theta \Gamma]^{-1} x_t
\]

Anticipated Supply Shocks

We also want to understand the dynamics following the announcement at time \( t = T \) of a future supply shock. To mimic the announcement of a future supply shock, we assume that \( q_t \) jumps up at \( t = T \), but also increase the lagged demands of slow-moving investors so that the active float isunchange. Specifically, letting \( \varepsilon_t [X_t] = E_t [X_{t+j}] - E_{t-1} [X_{t+j}] \) for some \( j \geq 0 \) denote the innovation to some random process, the supply shock is such that

\[
0 = \varepsilon_T [q_T] - (1-p) k^{-1} \sum_{j=1}^{k-1} \varepsilon_T [d_{T-j}] = [q_1 e_s - (1-p)k^{-1} 1_{(d)}] \varepsilon_T [X_T]
\]

For simplicity, we assume there is no change in expectation of future short rates or defaults (Formally, this means \( e_r^\prime [I - \theta \Gamma]^{-1} \varepsilon_T [X_T] = e_s^\prime [I - \theta \Gamma]^{-1} \varepsilon_T [X_T] = 0 \).)

Note that we can also allow for different announced paths of purchases. Specifically, by holding \( \varepsilon_T [q_T] \) fixed, we can vary the path of announced supply shocks by varying \( \{ \varepsilon_T [d_{T-j}] \} \) so long as we hold \( \sum_{j=1}^{k-1} \varepsilon_T [d_{T-j}] = (k/ (1-p)) \varepsilon_T [d_{T-j}] \). By holding \( \varepsilon_T [q_T] \) and \( \sum_{j=1}^{k-1} \varepsilon_T [d_{T-j}] \) fixed, we hold fixed the cumulative supply shock, but simply alter the announced timing of the shock.
Exercises like this can be used to evaluate different strategies for LSAP communications (i.e., announcement of purchases or tapering). Specifically, holding fixed the size of the purchase, we can ask whether we get a larger bang for buck using a gradual purchase strategy, a rapid purchase strategy, etc.

Upon announcement, the innovation to expected returns is

\[
\varepsilon_T [r x_{T+1}^L] = -\left(\frac{\theta}{1-\theta} a_r\right)^2 \sigma_r^2 + \left(\frac{\theta}{1-\theta} a_s\right)^2 \sigma_s^2 + \left(\frac{\theta}{1-\theta} a_z + (1-\theta)\right)^2 \sigma_z^2 \tau_{PB} (1-p)k^{-1}\delta_1 \varepsilon_T [x_T]
\]

So an anticipated shock only changes 1-period excess returns today to the extent that it induces a change in the demand of slow-moving investors (i.e., \(\delta_1 \varepsilon_T [x_T] \neq 0\)). However, for \(j > 1\), we have

\[
\varepsilon_T [r x_{T+j}^L] = \left(\frac{\theta}{1-\theta} a_r\right)^2 \sigma_r^2 + \left(\frac{\theta}{1-\theta} a_s\right)^2 \sigma_s^2 + \left(\frac{\theta}{1-\theta} a_z + (1-\theta)\right)^2 \sigma_z^2 \tau_{PB} [q_1 e_s - (1-p)k^{-1}(1(d) + \delta_1)]' \varepsilon_T [x_{T+j}]
\]

The innovation to yields is given by

\[
\varepsilon_T [y_T] = \left(\frac{\theta}{1-\theta} a_r\right)^2 \sigma_r^2 + \left(\frac{\theta}{1-\theta} a_s\right)^2 \sigma_s^2 + \left(\frac{\theta}{1-\theta} a_z + (1-\theta)\right)^2 \sigma_z^2 \tau_{PB} \times (1-\theta) \sum_{i=0}^{\infty} (\theta)^i \varepsilon_T [b_{T+i}^L]
\]

\[
= \left(\frac{\theta}{1-\theta} a_r\right)^2 \sigma_r^2 + \left(\frac{\theta}{1-\theta} a_s\right)^2 \sigma_s^2 + \left(\frac{\theta}{1-\theta} a_z + (1-\theta)\right)^2 \sigma_z^2 \tau_{PB} [q_1 e_s - (1-p)k^{-1}(1(d) + \delta_1)]' (1-\theta) [I - \theta \Gamma]^{-1} \varepsilon_T [x_T]
\]

Thus, we are going to see a significant jump up in yields at \(t = T\) because of the innovation to expectations (\(\varepsilon_T [x_T]\)) raises future expected term premia and hence perpetuity yields today. Further, even in the absence of additional innovations in expectations, perpetuity yields will follow a non-trivial path following the shock because the highest values of \(E_T [rx_{T+i}]\) will get a larger weight as they become closer to the present.

C.2 Two Asset Model

We now develop a two-asset version of the model. The model rests on the central premise of limited arbitrage theory that the marginal investor in any asset is a specialist with a large undiversified exposure to that asset. However, following Duffie (2010), we assume that there are a group of slow-moving investors who attempt to integrate these segmented markets at longer horizons. As a result, market will be more segmented in the short-run and more integrated in the long run. However, to the extent that integrating these markets is risky and the risk-tolerance of generalists is limited, these segmented markets will not be perfectly integrated even in the long-run.

This simple perspective on assets markets yield a rich range of predictions about asset pricing dynamics and the connections between dynamics in distinct markets.
C.2.1 Model Set Up

Assets

There are two long-term fixed income assets. The first asset, denoted $A$, is default–free and is only exposed to interest rate risk. Thus, we have

$$r_{t+1}^A = \frac{1}{1 - \theta_A} y_t^A - \frac{\theta_A}{1 - \theta_A} y_{t+1}^A$$  \hspace{1cm} (60)$$

where $\theta_A = 1/(1 + C_A)$. By contrast, the second asset, denoted $B$, has random cashflows each period. Thus, asset $B$ is exposed to cashflow risk as well as interest rate risk. Specifically, we have

$$r_{t+1}^B = \frac{1}{1 - \theta_B} y_t^B - \frac{\theta_B}{1 - \theta_B} y_{t+1}^B - (1 - \theta_B) z_{t+1}$$  \hspace{1cm} (61)$$

where $\theta_B = 1/(1 + C_B)$. The process for the short rate $r_t$ and the cash flow process $z_t$ are the same as above.

However, we now allow these two markets to be subject to different supply shocks. Specifically, the net supply that bond investors must hold in $A$ is

$$q_t^A = q_0^A + q_1^A s_t^A$$  \hspace{1cm} (62)$$

where

$$s_{t+1}^A = \rho_s^A s_t^A + \varepsilon_{s,t+1}^A$$  \hspace{1cm} (63)$$

Similarly, the net supply that bond investors must hold in $B$ is

$$q_t^B = q_0^B + q_1^B s_t^B$$  \hspace{1cm} (64)$$

where

$$s_{t+1}^B = \rho_s^B s_t^B + \varepsilon_{s,t+1}^B$$  \hspace{1cm} (65)$$

As above, it will be convenient to assume that $\varepsilon_{r,t+1}, \varepsilon_{z,t+1}, \varepsilon_{s,t+1}^A, \varepsilon_{s,t+1}^B$ are mutually orthogonal. However, it is straightforward to relax this assume.

Market Participants

There is a unit measure of agents, each with risk tolerance $\tau$. There are three types of agents who are distinguished by their ability to transact in different markets and by the frequency with which the can rebalance their portfolios. Fast-moving specialists in $A$ allocate between short-rate and long-term asset in market $A$ each period. Fast-moving $A$ specialists are present in mass $p_A$ and we let $b_t^A$ denote their demand for asset $A$. Fast-moving specialists in $B$ allocate between short-rate and long-term asset in the $B$ market each period. Fast-moving $B$ specialists are present in mass $p_B$ and we let $b_t^B$ denote their demand for asset $B$. Finally, there is a group of slow-moving generalists who allocate between both markets every $k$ periods. Slow-moving generalists are present in mass $1 - p_A - p_B$ and we let $d_t^A$ and $d_t^B$ denote the demand of the generalists who are active at time $t$ in market $A$ and $B$, respectively. Thus, by varying $p_A$ and $p_B$ we ask how price dynamics change as we vary the mix of specialists and generalists, holding constant total investor risk tolerance.
Equilibrium Conjecture

Suppose, for example, that $k = 5$. We conjecture an equilibrium involving the state vector

$$
x_t = \begin{bmatrix}
  r_t - \bar{r} \\
  s_t^A \\
  d_{t-1}^A - \delta_0^A \\
  d_{t-2}^A - \delta_0^A \\
  d_{t-3}^A - \delta_0^A \\
  d_{t-4}^A - \delta_0^A \\
  z_t - \bar{z} \\
  s_t^B \\
  d_{t-1}^B - \delta_0^B \\
  d_{t-2}^B - \delta_0^B \\
  d_{t-3}^B - \delta_0^B \\
  d_{t-4}^B - \delta_0^B
\end{bmatrix}.
$$

(66)

We conjecture that long-term yields in market $A$ and $B$ are

$$
y_t^A = \alpha_{A0} + \alpha_{A1}^t x_t
$$

(67)

$$
y_t^B = \alpha_{B0} + \alpha_{B1}^t x_t.
$$

(68)

and that the demands of slow-moving generalists are of the form

$$
d_t^A = \delta_{A0} + \delta_{A1}^t x_t
$$

(69)

$$
d_t^B = \delta_{B0} + \delta_{B1}^t x_t
$$

(70)

These assumptions imply that the state vector follows an AR(1) process. Critically, the transition matrix $\Gamma$ is a function of inattentive demand so we write $\Gamma = \Gamma (\delta_1^A, \delta_1^B)$. Specifically, we
have

\[ \mathbf{x}_{t+1} = \Gamma(\delta_{1A}, \delta_{1B}) \mathbf{x}_t + \epsilon_{t+1} \]

\[
\begin{bmatrix}
\rho_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \rho^A_s & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\delta^A_1 & \delta^A_2 & \delta^A_3 & \delta^A_4 & \delta^A_5 & \delta^A_6 & \delta^A_7 & \delta^A_8 & \delta^A_9 & \delta^A_{10} & \delta^A_{11} & \delta^A_{12} \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\delta^B_1 & \delta^B_2 & \delta^B_3 & \delta^B_4 & \delta^B_5 & \delta^B_6 & \delta^B_7 & \delta^B_8 & \delta^B_9 & \delta^B_{10} & \delta^B_{11} & \delta^B_{12} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
\end{bmatrix}
\begin{bmatrix}
r_t - \bar{r} \\
s^A_t \\
d^A_{t-1} - \delta^A_0 \\
d^A_{t-2} - \delta^A_0 \\
d^A_{t-3} - \delta^A_0 \\
d^A_{t-4} - \delta^A_0 \\
\end{bmatrix}
+ \begin{bmatrix}
r_{t+1} \\
s^A_{t+1} \\
d^A_{t+1} - \delta^A_0 \\
d^A_{t+1} - \delta^A_0 \\
d^A_{t+1} - \delta^A_0 \\
d^A_{t+1} - \delta^A_0 \\
\end{bmatrix}
\begin{bmatrix}
\epsilon_{r,t+1} \\
\epsilon^A_{s,t+1} \\
\epsilon_{z,t+1} \\
\epsilon^B_{s,t+1} \\
\end{bmatrix}
\]

where

\[
\text{Var}[\epsilon_{t+1}] = \Sigma = \begin{bmatrix}
\sigma^2_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \sigma^2_{sA} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

Fast-Moving A Specialists

Under our conjecture that \( y^A_t = \alpha_{A0} + \alpha^A_{A1} x_t \), we have

\[
rx^A_{t+1} = \frac{1}{1 - \theta_A} y^A_t - \frac{\theta_A}{1 - \theta_A} y^A_{t+1} - r_t \\
= (\alpha_{A0} - \bar{r}) + \left( \frac{1}{1 - \theta_A} \alpha_{A1} - e_r \right) x_t - \left( \frac{\theta_A}{1 - \theta_A} \alpha_{A1} \right) x_{t+1}
\]
and

\[
E_t \left[ r x^A_{t+1} \right] = (\alpha_{A0} - \tau) + \left( \frac{1}{1 - \theta_A} \alpha_{A1} - e_r \right)' x_t - \left( \frac{\theta_A}{1 - \theta_A} \alpha_{A1} \right)' \Gamma x_t
\]

\[
Var_t \left[ r x^A_{t+1} \right] = \left( \frac{\theta_A}{1 - \theta_A} \right)^2 Var_t \left[ y^A_{t+1} \right]
\]

\[
= \left( \frac{\theta_A}{1 - \theta_A} \alpha_{A1} \right)' \Sigma \left( \frac{\theta_A}{1 - \theta_A} \alpha_{A1} \right)
\]

\[
= \left( \frac{\theta_A}{1 - \theta_A} a_{Ar} \right)^2 \sigma_r^2 + \left( \frac{\theta_A}{1 - \theta_A} a_{As} \right)^2 \sigma_s^2
\]

Specialists solve

\[
\max_{b_t} \left\{ E_t \left[ r x^A_{t+1} \right] b_t^A - \frac{(b_t^A)^2}{2\tau} Var_t \left[ r x^A_{t+1} \right] \right\}
\]

so we have

\[
b_t^A = \tau E_t \left[ r x^A_{t+1} \right] = \left[ \tau - \left( \frac{\theta_A}{1 - \theta_A} a_{Ar} \right)^2 \sigma_r^2 + \left( \frac{\theta_A}{1 - \theta_A} a_{As} \right)^2 \sigma_s^2 \right] + \left[ \tau \left( \frac{\theta_A}{1 - \theta_A} \alpha_{A1} - e_r \right)' - \frac{\theta_A}{1 - \theta_A} \alpha_{A1}' \Gamma \right] x_t
\]

\[
\text{Fast-Moving } B \text{ Specialists}
\]

Under our conjecture that \( y^B_t = \alpha_{B0} + \alpha_{B1} x_t \), we have

\[
rx^B_{t+1} = y^B_t - t - \frac{\theta_B}{1 - \theta_B} (y^B_{t+1} - y^B_t) - (1 - \theta_B) z_{t+1}
\]

\[
= \alpha_{B0} - \tau - (1 - \theta_B) z + (\alpha_{B1} - e_r)' x_t - \frac{\theta_B}{1 - \theta_B} (\alpha_{B1}' x_{t+1} - \alpha_{B1} x_t) - (1 - \theta_B) e_z' x_{t+1}
\]

\[
= (\alpha_{B0} - \tau - (1 - \theta_B) z) + \left( \frac{1}{1 - \theta_B} \alpha_{B1} - e_r \right)' x_t - \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right)' x_{t+1}
\]

and

\[
E_t \left[ rx^B_{t+1} \right] = (\alpha_{B0} - \tau - (1 - \theta_B) z) + \left( \frac{1}{1 - \theta_B} \alpha_{B1} - e_r \right)' x_t - \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right)' \Gamma x_t
\]

\[
Var_t \left[ rx^B_{t+1} \right] = \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right)' \Sigma \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right)
\]

\[
= \left( \frac{\theta_B}{1 - \theta_B} a_{Br} \right)^2 \sigma_r^2 + \left( \frac{\theta_B}{1 - \theta_B} a_{Bs} \right)^2 \sigma_s^2 + \left( \frac{\theta_B}{1 - \theta_B} a_{Bz} + (1 - \theta_B) \right)^2 \sigma_z^2
\]

Specialists solve

\[
\max_{b_t} \left\{ E_t \left[ rx^B_{t+1} \right] b_t^B - \frac{(b_t^B)^2}{2\tau} Var_t \left[ rx^B_{t+1} \right] \right\}
\]
so we have

\[
b_t^B = \frac{E_t \left[ r_{x_{t+1}^B} \right]}{\text{Var} \left[ r_{x_{t+1}^B} \right]}
\]

\[
= \tau \left[ \alpha_{B0} - \tau - (1 - \theta_B) \bar{z} \right] + \left[ \left( \frac{\theta_B}{1 - \theta_B} a_{Br} \right)^2 \sigma_r^2 + \left( \frac{\theta_B}{1 - \theta_B} a_{Bx} \right)^2 \sigma_s^2 + \left( \frac{\theta_B}{1 - \theta_B} a_{Bz} + (1 - \theta_B) \right)^2 \sigma_z^2 \right]
\]

\[
+ \left[ \left( 1 - \theta_B a_{B1} - e_r \right)' - \left( \frac{\theta_B}{1 - \theta_B} a_{B1} + (1 - \theta_B) e_z \right)' \Gamma \right] \mathbf{x}_t
\]

**Slow-Moving Generalists**

Slow-moving generalists solve

\[
\max_{d_t^A, d_t^B} \left\{ \begin{array}{c}
d_t^A E_t \left[ \sum_{i=1}^k r_{x_{t+i}^A} \right] + d_t^B E_t \left[ \sum_{i=1}^k r_{x_{t+i}^B} \right] + E_t \left[ \sum_{i=0}^{k-1} r_{t+i} \right] - \frac{1}{2\tau} \left( d_t^A \right)^T V_{ar}^A \left[ \sum_{i=1}^k r_{x_{t+i}^A} \right] + (d_t^B)^T V_{ar}^B \left[ \sum_{i=1}^k r_{x_{t+i}^B} \right] - 2d_t^A E_t \left[ \sum_{i=1}^k r_{x_{t+i}^A} \right] \left[ \sum_{i=1}^k r_{x_{t+i}^B} \right] + 2d_t^B E_t \left[ \sum_{i=1}^k r_{x_{t+i}^A} \right] \left[ \sum_{i=1}^k r_{x_{t+i}^B} \right] \right. \\
\left. + 2d_t^A d_t^B \text{Cov}_t \left[ \sum_{i=1}^k r_{x_{t+i}^A} \left[ \sum_{i=1}^k r_{x_{t+i}^B} \right] + 2d_t^B \text{Cov}_t \left[ \sum_{i=1}^k r_{x_{t+i}^A} \left[ \sum_{i=1}^k r_{x_{t+i}^B} \right] \right] \right. \\
\right. \end{array} \right\}
\]

which implies

\[
\begin{bmatrix}
d_t^A \\
d_t^B
\end{bmatrix} = \left[ \delta_{A0} + \delta_{A1}' \mathbf{x}_t \right]
\]

\[
= \tau \left[ \begin{array}{cc}
V_{Ak} & C_{ABk} \\
C_{ABk} & V_{Bk}
\end{array} \right]^{-1} \left( \begin{array}{c}
E_t \left[ \sum_{i=1}^k r_{x_{t+i}^A} \right] \\
E_t \left[ \sum_{i=1}^k r_{x_{t+i}^B} \right]
\end{array} \right) - \left( \begin{array}{c}
\text{Cov}_t \left[ \sum_{i=1}^k r_{x_{t+i}^A} \right] \\
\text{Cov}_t \left[ \sum_{i=1}^k r_{x_{t+i}^B} \right]
\end{array} \right) \left( \begin{array}{c}
\sum_{i=1}^k r_{x_{t+i}^A} \left[ \sum_{i=1}^k r_{x_{t+i}^B} \right] \\
\sum_{i=1}^k r_{x_{t+i}^A} \left[ \sum_{i=1}^k r_{x_{t+i}^B} \right]
\end{array} \right)
\]

\[
= \frac{\tau}{V_{Ak} V_{Bk} - (C_{ABk})^2} \left( \begin{array}{c}
V_{Bk} E_t \left[ \sum_{i=1}^k r_{x_{t+i}^A} \right] - C_{ABk} E_t \left[ \sum_{i=1}^k r_{x_{t+i}^B} \right] \\
V_{Ak} E_t \left[ \sum_{i=1}^k r_{x_{t+i}^B} \right] - C_{ABk} E_t \left[ \sum_{i=1}^k r_{x_{t+i}^A} \right]
\end{array} \right)
\]

\[
= \frac{\tau}{V_{Ak} V_{Bk} - (C_{ABk})^2} \left( V_{Bk} C_{Ak} - C_{ABk} C_{Bk} \right)
\]

Realized returns are

\[
\sum_{j=1}^k r_{x_{t+1}^A} = k (\alpha_{A0} - \tau) + (\alpha_{A1} - e_r)' \left( \sum_{j=0}^{k-1} x_{t+j} \right) - \frac{\theta_A}{1 - \theta_A} \alpha_{A1}' (x_{t+k} - x_t)
\]

\[
\sum_{j=1}^k r_{x_{t+1}^B} = k (\alpha_{B0} - \tau - (1 - \theta_B) \bar{z}) + (\alpha_{B1} - e_r)' \left( \sum_{j=0}^{k-1} x_{t+j} \right) - (1 - \theta_B) e_z' \left( \sum_{j=1}^k x_{t+j} \right) - \frac{\theta_B}{1 - \theta_B} (\alpha_{B1}' (x_{t+k} - x_t - \alpha_{B1}' x_t))
\]
Expected returns are

\[
E_t \left[ \sum_{j=1}^{k} r_{t+j}^A \right] = k (\alpha_{A0} - \tau) + (\alpha_{A1} - e_r)' (I - \Gamma)^{-1} \left( I - \Gamma^k \right) x_t - \left( \frac{\theta_A}{1 - \theta_A} \right) \alpha_{A1} \left( \Gamma^j - I \right) x_t
\]

\[
E_t \left[ \sum_{j=1}^{k} r_{t+j}^B \right] = k (\alpha_{B0} - \tau - (1 - \theta_B) \bar{z}) + (\alpha_{B1} - e_r)' (I - \Gamma)^{-1} \left( I - \Gamma^k \right) x_t
\]

\[
- (1 - \theta_B) e_r' (I - \Gamma)^{-1} \left( \Gamma^k - I \right) x_t - \left( \frac{\theta_B}{1 - \theta_B} \right) \alpha_{B1} \left( \Gamma^k - I \right) x_t
\]

The variance of $A$ market returns is

\[
V_{A_k} (\alpha, \delta) = \text{Var}_t \left[ (\alpha_{A1} - e_r)' \left( \sum_{j=1}^{k-1} x_{t+j} \right) - \left( \frac{\theta_A}{1 - \theta_A} \right) \alpha_{A1} x_{t+k} \right] = (\alpha_{A1} - e_r)' \left( \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} C_{[t+i,t+j]} \right) (\alpha_{A1} - e_r) + \left( \frac{\theta_A}{1 - \theta_A} \right)^2 \alpha_{A1} C_{[t+k,t+k]} \alpha_{A1}
\]

\[
- 2 \left( \frac{\theta_A}{1 - \theta_A} \right) (\alpha_{A1} - e_r)' \sum_{i=1}^{k-1} C_{[t+i,t+k]} \alpha_{A1}
\]

The variance of $B$ market returns is

\[
V_{B_k} (\alpha, \delta) = \text{Var}_t \left[ (\alpha_{B1} - e_r - (1 - \theta_B) e_z)' \left( \sum_{j=1}^{k-1} x_{t+j} \right) - \left( \frac{\theta_B}{1 - \theta_B} \right) \alpha_{B1} (1 - \theta_B) e_z \right] \left( \sum_{j=1}^{k-1} x_{t+j} \right) - \left( \frac{\theta_B}{1 - \theta_B} \right) \alpha_{B1} + (1 - \theta_B) e_z \right] x_{t+k} = (\alpha_{B1} - e_r - (1 - \theta_B) e_z)' \left( \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} C_{[t+i,t+j]} \right) (\alpha_{B1} - e_r - (1 - \theta_B) e_z)
\]

\[
+ \left( \frac{\theta_B}{1 - \theta_B} \right) \alpha_{B1} + (1 - \theta_B) e_z \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right)
\]

\[
- 2 (\alpha_{B1} - e_r - (1 - \theta_B) e_z)' \sum_{i=1}^{k-1} C_{[t+i,t+k]} \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right)
\]

The covariance between $A$ and $B$ market returns is

\[
C_{A_{Bk}} (\alpha, \delta) = \text{Cov}_t \left[ (\alpha_{A1} - e_r)' \left( \sum_{j=1}^{k-1} x_{t+j} \right) - \left( \frac{\theta_A}{1 - \theta_A} \right) \alpha_{A1} x_{t+k}, (\alpha_{B1} - e_r - (1 - \theta_B) e_z)' \left( \sum_{j=1}^{k-1} x_{t+j} \right) - \left( \frac{\theta_B}{1 - \theta_B} \right) \alpha_{B1} + (1 - \theta_B) e_z \right] x_{t+k} = (\alpha_{A1} - e_r)' \left( \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} C_{[t+i,t+j]} \right) (\alpha_{B1} - e_r - (1 - \theta_B) e_z)
\]

\[
- (\alpha_{A1} - e_r)' \sum_{i=1}^{k-1} C_{[t+i,t+k]} \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right)
\]

\[
- \left( \frac{\theta_A}{1 - \theta_A} \right) \alpha_{A1} \sum_{i=1}^{k-1} C_{[t+i,t+k]} \left( \alpha_{B1} - e_r - (1 - \theta_B) e_z \right)
\]

\[
+ \left( \frac{\theta_A}{1 - \theta_A} \right) \alpha_{A1} C_{[t+k,t+k]} \left( \alpha_{B1} - e_r - (1 - \theta_B) e_z \right)
\]

Finally, the covariance of $A$ market returns with the return on cash is

\[
C_{A_k} = \text{Cov}_t \left[ (\alpha_{A1} - e_r)' \left( \sum_{j=1}^{k-1} x_{t+j} \right) - \left( \frac{\theta_A}{1 - \theta_A} \right) \alpha_{A1} x_{t+k}, e_r' \left( \sum_{j=1}^{k-1} x_{t+j} \right) \right] = (\alpha_{A1} - e_r)' \left( \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} C_{[t+i,t+j]} \right) e_r
\]

\[
- \left( \frac{\theta_A}{1 - \theta_A} \right) \alpha_{A1} \sum_{i=1}^{k-1} C_{[t+i,t+k]} e_r
\]
Similarly, for market $B$

$$C_{Bk} = \text{Cov}_t \left[ (\alpha_{B1} - e_r - (1 - \theta_B) e_z)' \left( \sum_{j=1}^{k-1} x_{t+j} \right) - \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right)' x_{t+k} e_r' \left( \sum_{j=1}^{k-1} x_{t+j} \right) \right]$$

$$= (\alpha_{B1} - e_r - (1 - \theta_B) e_z)' \left( \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} C_{[i+j]} \right) e_r$$

$$- \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right)' \sum_{i=1}^{k-1} C_{[i+j+k]} e_r$$

Solution for Market A

Market clearing: demand (LHS) = supply (RHS)

$$(1 - p_A - p_B) k^{-1} d_t^A + p_A b_t^A = q_t^A - (1 - p_A - p_B) \left( k^{-1} \sum_{i=1}^{k-1} d_{t-i}^A \right)$$

Demand

$$(1 - p_A - p_B) k^{-1} d_t^A + p_A b_t^A = \left[ (1 - p_A - p_B) k^{-1} \delta_{A0} + p_A \tau \frac{(\alpha_{A0} - \tau)}{(D_A - 1)^2 \alpha'_{A1} \Sigma \alpha_{A1}} \right]$$

$$+ \left[ (1 - p_A - p_B) k^{-1} \delta'_{A1} + p_A \tau \frac{(\alpha_{A1} - e_r)' - (D_A - 1) \alpha'_{A1} (I - \Gamma)}{(D_A - 1)^2 \alpha'_{A1} \Sigma \alpha_{A1}} \right] x_t$$

Supply

$$q_t^A - (1 - p_A - p_B) \left( k^{-1} \sum_{i=1}^{k-1} d_{t-i}^A \right)$$

$$= q_0^A + q_1^A s_t^A - (1 - p_A - p_B) \frac{(k-1)}{k} \delta_0^A - (1 - p_A - p_B) k^{-1} \sum_{i=1}^{k-1} (q_{t-i}^A - \delta_0^A)$$

$$= \left( q_0^A - (1 - p_A - p_B) \frac{(k-1)}{k} \delta_0^A \right) + (q_1^A e_{s,A} - (1 - p_A - p_B) k^{-1} I_{(A)})' x_t$$

where $I_{(A)} = e_3 + e_4 + e_5 + e_6$.

Matching constants, we have

$$\alpha_{A0} \tau + \frac{(D_A - 1)^2 \alpha'_{A1} \Sigma \alpha_{A1}}{p_A} (q_0^A - (1 - p_A - p_B) \delta_{A0})$$

Matching slope coefficients, we have

$$\alpha_{A1} = (1 - \rho) \left[ I - \rho \Gamma' \right]^{-1} e_r$$

$$+ \frac{(D_A - 1)^2 \alpha'_{A1} \Sigma \alpha_{A1}}{p_A \tau} (1 - \rho) \left[ I - \rho \Gamma' \right]^{-1} \left[ q_1^A e_{s,A} - (1 - p_A - p_B) k^{-1} I_{(A)} + \delta_{A1} \right]$$

Solve for Market B
We have

\[
p_B \tau = \frac{\alpha_{B0} - \tau - (1 - \theta_B) z}{\left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right)'} \Sigma \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right)
\]

\[+ p_B \left[ \tau \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} - e_r \right) - \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right) \right] x_t \]

\[= (q_0^B - (1 - p_A - p_B) \delta_0^B) + (q_1^B e_{sB} - (1 - p_A - p_B) k^{-1} (1_{(A)} + \delta_1 B))' x_t \]

Matching constants, we have

\[
\alpha_{B0} = \tau + (1 - \theta_B) z
\]

\[+ \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right) \Sigma \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right) (q_0^B - (1 - p_A - p_B) \delta_0^B) \]

Matching slope coefficients, we have

\[
\alpha_{B1} = (1 - \theta_B) \left[ I - \theta_B \Gamma' \right]^{-1} \left[ e_r + (1 - \theta_B) \Gamma' e_z \right]
\]

\[+ \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right) \Sigma \left( \frac{\theta_B}{1 - \theta_B} \alpha_{B1} + (1 - \theta_B) e_z \right) (1 - \theta_B) \left[ I - \theta_B \Gamma' \right]^{-1} \left[ q_1^B e_{sB} - (1 - p_A - p_B) k^{-1} (1_{(A)} + \delta_1 B) \right] \]

C.2.3 Equilibrium Yields and Excess Returns

Yields are

\[y_t^A = \alpha_{A0} + \alpha_{A1}' x_t\]

Expectations of future short rates

\[= \tau + (1 - \theta_A) e_r' \left[ I - \theta_A \Gamma' \right]^{-1} x_t\]

Uncond Term Premia Contrib.

\[+ \left( \frac{\theta_A}{1 - \theta_A} a_{Ar} \right)^2 \sigma_r^2 + \left( \frac{\theta_A}{1 - \theta_A} a_{As} \right)^2 \sigma_s^2 \]

\[\frac{p_A}{\tau p_A} \left( q_0^A - (1 - p_A - p_B) \delta_{A0} \right) \]

Cond Term Premium Contrib.

\[+ \left[ \left( \frac{\theta_A}{1 - \theta_A} a_{Ar} \right)^2 \sigma_r^2 + \left( \frac{\theta_A}{1 - \theta_A} a_{As} \right)^2 \sigma_s^2 \right] \left[ q_1^A e_s - (1 - p_A - p_B) k^{-1} (1_{(A)} + \delta_{A1}) \right]' (1 - \theta_A) \left[ I - \theta_A \Gamma' \right]^{-1} x_t \]
and

\[ y^B_t = \alpha^B_0 + \alpha^B_1 x_t \]

**Expectations of future short rates**

\[ \tau + (1 - \theta^B) e^0_t [I - \theta^B \Gamma]^{-1} x_t + (1 - \theta^B) \left( \tau + (1 - \theta^B) e^0_t [I - \theta^B \Gamma]^{-1} x_t \right) \]

**Uncond Term Premia Contrib.**

\[ \frac{2 \sigma^2_r}{\tau \theta^B} \left( \theta^B (1 - \theta^B) a^B_r \right) + \frac{2 \sigma^2_s}{\tau \theta^B} \left( \theta^B (1 - \theta^B) a^B_s \right) + \frac{2 \sigma^2_{sA}}{\tau \theta^B} \left( \theta^B (1 - \theta^B) a^A_s \right) \]

**Cond Term Premium Contrib.**

\[ \frac{2 \sigma^2_r}{\tau \theta^B} \left( \theta^B (1 - \theta^B) a^B_r \right) + \frac{2 \sigma^2_s}{\tau \theta^B} \left( \theta^B (1 - \theta^B) a^B_s \right) + \frac{2 \sigma^2_{sA}}{\tau \theta^B} \left( \theta^B (1 - \theta^B) a^A_s \right) \]

**Expectations of future defaults**

\[ \mathbb{E}[\text{default}_t] = \left( q^A_0 - (1 - p^A - p^B) \delta^B_0 \right) \]

**Expected 1-period excess returns are**

\[ E_t [r^A_{t+1}] = \frac{2 \sigma^2_r}{\tau \theta^A} \left( \theta^A (1 - \theta^A) a^A_r \right) + \frac{2 \sigma^2_s}{\tau \theta^A} \left( \theta^A (1 - \theta^A) a^A_s \right) \]

\[ \left( q^A_0 - (1 - p^A - p^B) \delta^A_0 \right) \]

and

\[ E_t [r^B_{t+1}] = \frac{2 \sigma^2_r}{\tau \theta^B} \left( \theta^B (1 - \theta^B) a^B_r \right) + \frac{2 \sigma^2_s}{\tau \theta^B} \left( \theta^B (1 - \theta^B) a^B_s \right) + \frac{2 \sigma^2_{sA}}{\tau \theta^B} \left( \theta^B (1 - \theta^B) a^A_s \right) \]

\[ \left( q^B_0 - (1 - p^A - p^B) \delta^B_0 \right) \]

\[ \left[ q^B_1 e^B_s - (1 - p^A - p^B) k^{-1} (1 + (\delta^B_1)) \right] x_t \]