Liquidity Risk and the
Dynamics of Arbitrage Capital

Péter Kondor

Central European University and CEPR

Dimitri Vayanos

London School of Economics, CEPR and NBER

June 20, 2014[*]

Abstract

We develop a dynamic model of liquidity provision, in which hedgers can trade multiple risky assets with arbitrageurs. We compute the equilibrium in closed form when arbitrageurs' utility over consumption is logarithmic or risk-neutral with a non-negativity constraint. Liquidity is increasing in arbitrageur wealth, while asset volatilities, correlations, and expected returns are hump-shaped. Liquidity is a priced risk factor: assets that suffer the most when liquidity decreases, e.g., those with volatile cashflows or in high supply by hedgers, offer the highest expected returns. When hedging needs are strong, arbitrageurs can choose to provide less liquidity even though liquidity provision is more profitable.

1 Introduction

Liquidity in financial markets is mostly provided by specialized agents, such as dealers, brokers and hedge funds; all partially playing the role of market makers and arbitrageurs. There is increasing amount of evidence that the fluctuation of capital of these institutions is closely linked to asset prices: while past returns affect their level of capital, the capital of these agents also affects returns.[1]

In this paper we build a parsimonious, analytically tractable, dynamic equilibrium model to understand the interaction between asset prices, liquidity and these special agents' capital. Our model rationalize why liquidity varies over time and in a correlated manner across assets, why assets differ in their covariance with aggregate liquidity (i.e., their liquidity beta), and why that covariance is linked to expected returns in the cross-section.[2] In addition, our model provides a broader framework for analyzing the risk-sharing across heterogeneous agents, the dynamics of their capital and its link with asset prices. In important special cases, our framework admits closed form expressions for all equilibrium objects.

We assume a continuous-time infinite-horizon economy with two sets of competitive agents: hedgers, who receive a risky income flow, and arbitrageurs (our collective term for specialized agents providing liquidity), who can absorb part of that risk in exchange for compensation. Agents can invest in a riskless linear technology, and in multiple risky assets whose prices are endogenously determined in equilibrium. Arbitrageurs provide liquidity to hedgers by taking the other side of their trades, as well as insurance by absorbing their risk. In our baseline specification, hedgers' risk aversion and income variance are constant over time, implying a constant demand for liquidity. Arbitrageurs are infinitely lived and have constant relative risk aversion (CRRA) utility over intertemporal consumption. This implies that the supply of liquidity is time-varying because arbitrageurs' absolute risk aversion depends on their wealth.

Our starting point is the observation that the portfolio of the representative arbitrageur must

---

[1]For example, Comerton-Forde, Hendershott, Jones, Moulton, and Seasholes (2010) find that bid-ask spreads quoted by specialists in the New York Stock Exchange widen when specialists experience losses. Jylha and Suominen (2011) find that outflows from hedge funds that perform the carry trade predict poor performance of that trade, with low interest-rate currencies appreciating and high-interest rate ones depreciating. Froot and O'Connell (1999) find that following losses by catastrophe insurers, premia increase even for risks unrelated to the losses.

[2]There is a large and increasing empirical literature characterizing assets' liquidity and searching for priced liquidity factors. In particular, Chordia, Roll, and Subrahmanyam (2000), Hasbrouck and Seppi (2001), and Huberman and Halka (2001) find that liquidity varies over time and in a correlated manner across assets. Amihud (2002) and Hameed, Kang, and Viswanathan (2010) link time-variation in aggregate liquidity to the returns of the aggregate stock market. Pastor and Stambaugh (2003) and Acharya and Pedersen (2005) find that aggregate liquidity is a priced factor in the stock market. Sadka (2010) and Franzoni, Nowak, and Phalippou (2012) find that liquidity is a priced factor for hedge-fund and private-equity returns, respectively. For more references, see Vayanos and Wang (2013) who survey the theoretical and empirical literature on market liquidity.

be a pricing factor. It is so, because arbitrageurs first-order condition prices the assets. Because other investors have a hedging motive to trade, this porfolio is not the market portfolio. Assets that have high expected returns are the ones which covary highly with the representative arbitrageur's wealth, that is, with her portfolio.

While a factor based on arbitrageurs' portfolio is hard to measure, a liquidity beta, the covariance of the return of the asset with market-wide variation in liquidity (measured as price impact), captures the sensitivity of assets to this factor. For the argument, we define illiquidity based on the impact that hedgers have on prices, and show that it has a cross-sectional and a time-series dimension. In the cross-section, illiquidity is higher for assets with more volatile cashflows. In the time-series, illiquidity increases following losses by arbitrageurs. Hence, liquidity varies over time in response to changes in arbitrageur wealth, and this variation is common across assets. Because arbitrageurs sell a fraction of their portfolio following losses, assets that covary the most with that portfolio suffer the most when illiquidity increases. Therefore, the market-wide element in illiquidity is capturing the priced risk factor generated by arbitrageurs' portfolio. In particular, an asset's expected return is proportional (with a negative coefficient) to the covariance between its return and aggregate illiquidity. Other liquidity-related covariances used in empirical work, e.g., between an asset's illiquidity and aggregate illiquidity or return, are less informative about expected returns.

In addition to addressing liquidity risk, our model provides a broader framework to study risk-sharing across heterogeneous agents. We use this framework to derive a number of results on the dynamics of arbitrage capital and its link with asset prices and risk-sharing.

Equilibrium can be described by a single state variable, arbitrageur wealth. Wealth determines the arbitrageurs' effective risk aversion, which consists of two terms: one reflecting the myopic demand, and one reflecting the demand for intertemporal hedging. Even though the optimization of arbitrageurs is intertemporal, risk-sharing between them and hedgers follows a static optimal rule, with effective risk aversion replacing myopic risk aversion.

We derive closed-form solutions in two special cases of CRRA utility: logarithmic utility, and risk-neutrality with non-negative consumption. Effective risk aversion is the inverse of wealth in the logarithmic case, and an affine function of the cotangent of wealth in the risk-neutral case. When wealth increases, effective risk aversion decreases. As a consequence, arbitrageurs hold larger positions, and their profitability decreases.

Effective risk-aversion in the risk-neutral case is driven purely by the intertemporal hedging demand. Arbitrageurs take into account that in states where their portfolio performs poorly other

2

arbitrageurs also perform poorly and hence liquidity provision becomes more profitable. To have more wealth to provide liquidity in those states, arbitrageurs limit their investment in the risky assets, hence behaving as risk-averse. This behavior becomes more pronounced when hedgers become more risk averse or asset cashflows become more volatile. In both cases, arbitrageurs can choose to provide less liquidity even though liquidity provision becomes more profitable.

Because profitability is inversely related to arbitrageurs wealth level, arbitrageurs wealth follows a stationary distribution. In our special cases, this distribution can be characterized by a single parameter, which is increasing in hedger risk aversion and cashflow volatility. For small values of this parameter, the stationary distribution is either concentrated at zero or has a density that is decreasing in wealth. For larger values, the density becomes bimodal. The intuition for the bimodal shape is that when hedging needs are strong, liquidity provision is more profitable. Therefore, arbitrageur wealth grows fast, and large values of wealth can be more likely in steady state than small or intermediate values. At the same time, while profitability (per unit of wealth) is highest when wealth is small, wealth grows away from small values slowly in absolute terms. Therefore, small values of wealth are more likely than intermediate values.

Among other results, we also show that the feedback effects between arbitrageur wealth and asset prices, creates endogenous risk which takes the form of asset volatilities being hump-shaped in arbitrageur wealth. When wealth is small, shocks to wealth are small in absolute terms, and so is the price volatility that they generate. When wealth is large, arbitrageurs provide perfect liquidity to hedgers and prices are not sensitive to changes in wealth. Asset correlations and expected returns are similarly hump-shaped.

Finally, we relax the assumption that hedgers' risk aversion is constant over time by introducing infinitely lived hedgers with CARA utility. This implies that hedgers' effective risk aversion also varies with arbitrageurs' capital because of its effect on intertemporal hedging demand. We argue that this modification leaves are main insights intact.

To our knowledge, ours is the first paper to build an analytically tractable theory connecting liquidity risk to the capital of liquidity providers both in the cross-section and in the time-series. Still, our paper builds on a rich literature on liquidity and asset pricing.

A first group of related papers study the pricing of liquidity risk. In Holmstrom and Tirole (2001), illiquidity is defined in terms of firms' financial constraints. Firms avoid assets whose return is low when constraints are severe, and these assets offer high expected returns in equilibrium. Our result that arbitrageurs avoid assets whose return is low when liquidity provision becomes more

profitable has a similar flavor. The covariance between asset returns and illiquidity, however, is endogenous in our model because prices depend on arbitrageur wealth. In Amihud (2002) and Acharya and Pedersen (2005), illiquidity takes the form of exogenous time-varying transaction costs. An increase in the costs of trading an asset raises the expected return that investors require to hold it and lowers its price. A negative covariance between illiquidity and asset prices arises also in our model but because of an entirely different mechanism: high illiquidity and low prices are endogenous symptoms of low arbitrageur wealth. The endogenous variation in illiquidity is also what drives the cross-sectional relationship between expected returns and liquidity-related covariances. In contrast to Acharya and Pedersen (2005), we show that one of these covariances explains expected returns perfectly, while the others do not.

A second group of related papers link arbitrage capital to liquidity and asset prices. Some of these papers emphasize margin constraints. In Gromb and Vayanos (2002), arbitrageurs intermediate trade between investors in segmented markets, and are subject to margin constraints. Because of the constraints, the liquidity that arbitrageurs provide to investors increases in their wealth. In Brunnermeier and Pedersen (2009), margin-constrained arbitrageurs intermediate trade in multiple assets across time periods. Assets with more volatile cashflows are more sensitive to changes in arbitrageur wealth. Garleanu and Pedersen (2011) introduce margin constraints in an infinite-horizon setting with multiple assets. They show that assets with higher margin requirements earn higher expected returns and are more sensitive to changes in the wealth of the margin-constrained agents. This result is suggestive of a priced liquidity factor. In our model cross-sectional differences in assets' covariance with aggregate illiquidity arise because of differences in cashflow volatility and hedger supply rather than in margin constraints.

Other papers assume constraints on equity capital, which may be implicit (as in our paper) or explicit. In Xiong (2001) and Kyle and Xiong (2001), arbitrageurs with logarithmic utility over consumption can trade with long-term traders and noise traders over an infinite horizon. The liquidity that arbitrageurs can provide is increasing in their wealth, and asset volatilities are hump-shaped. In He and Krishnamurthy (2013), arbitrageurs can raise capital from other investors to invest in a risky asset over an infinite horizon, but this capital cannot exceed a fixed multiple of their internal capital. When arbitrageur wealth decreases, the constraint binds, and asset volatility and expected returns increase. In Brunnermeier and Sannikov (forthcoming), arbitrageurs are more efficient holders of productive capital. The long-run stationary distribution of their wealth can have a decreasing or a bimodal density. These papers mostly focus on the case of one risky asset (two assets in Kyle and Xiong (2001)), and hence cannot address the pricing of liquidity risk in the

cross-section.

Finally, our paper is related to the literature on consumption-based asset pricing with heterogeneous agents, e.g., Dumas (1989), Wang (1996), Chan and Kogan (2002), Bhamra and Uppal (2009), Garleanu and Panageas (2012), Basak and Pavlova (2013), Chabakauri (2013), Longstaff and Wang (2013). In these papers, agents have CRRA-type utility and differ in their risk aversion. As the wealth of the less risk-averse agents increases, Sharpe ratios decrease, and this causes volatilities to be hump-shaped. In contrast to these papers, we assume that only one set of agents has wealth-dependent risk aversion. This allows us to focus more sharply on the wealth effects of liquidity providers. We also fix the riskless rate through the linear technology, while in these papers the riskless rate is determined by aggregate consumption. Fixing the riskless rate simplifies our model and allows us to focus on price movements caused by changes in expected returns.

A methodological contribution relative to the above groups of papers is that we provide an analytically tractable model with multiple assets, dynamics, heterogeneous agents, and wealth effects. With a few exceptions, the dynamic models cited above compute the equilibrium by solving differential equations numerically. By contrast, we derive closed-form solutions, under both logarithmic and risk-neutral preferences, and prove analytically each of our main results.[3]

We proceed as follows. In Section 2 we present the model. In Section 3 we derive optimal risk-sharing, the price-of-risk and the wealth dynamics of hedgers and arbitrageurs by allowing only risky assets which are claims on the next instant cash-flow. We show that these results remain valid with long-lived assets in Section 4 where we also present the expressions for expected return and return correlation. In Section 5 we explore the implications of our model for liquidity risk. In Section 6 present our extension with infinitely lived CARA hedgers. Finally, we conclude.

2   Model

Time t is continuous and goes from zero to infinity. Uncertainty is described by the N-dimensional Brownian motion B_t. There is a riskless technology, whose instantaneous return is constant over time and equal to r. There are also N risky assets with the $N \times 1$ price vector S_t at time t and

---

[3]Closed-form solutions are also derived in Danielsson, Shin, and Zigrand (2012) and Gromb and Vayanos (2014). In the former paper, risk-neutral arbitrageurs are subject to a VaR constraint and can trade with long-term traders who submit exogenous demand functions. In the latter paper, arbitrageurs intermediate trade across segmented markets and are subject to margin constraints. Their activity involves no risk because the different legs of their trades cancel.

cash-flow $dD_{t'}$ for $t' \geq t$, where

$$dD_t = \bar{D}dt + \sigma^\top dB_t, \tag{2.1}$$

and $\bar{D}$ is a constant $N \times 1$ vector, σ is a constant and invertible $N \times N$ matrix, and $\top$ denotes transpose. The supply of these assets is s, an $N \times 1$ vector.

There are two sets of agents, hedgers and arbitrageurs. Each set forms a continuum with measure one. Hedgers choose asset positions at time t to maximize the mean-variance objective

$$\mathrm{E}_t(dv_t) - \frac{\alpha}{2}\mathrm{Var}_t(dv_t), \tag{2.2}$$

where dv_t is the change in wealth between t and $t + dt$, and α is a risk-aversion coefficient. To introduce hedging needs, we assume that hedgers receive a random endowment $u^\top dD_t$ at $t + dt$, where u is a constant $N \times 1$ vector. This endowment is added to dv_t. We set $\Sigma \equiv \sigma^\top \sigma$.

Since the hedgers' risk-aversion coefficient α and endowment variance $u^\top \Sigma u$ are constant over time, their demand for liquidity, derived in the next section, is also constant. We intentionally simplify the model in this respect, so that we can focus on the supply of liquidity, which is time-varying because of the wealth-dependent risk aversion of arbitrageurs.

One interpretation of the hedgers is as generations living over infinitesimal periods. The generation born at time t is endowed with initial wealth $\bar{v}$, and receives the additional endowment $u^\top dD_t$ at $t + dt$. It consumes all its wealth at $t + dt$ and dies. (In a discrete-time version of our model, each generation would be born in one period and die in the next.) If preferences over consumption are described by the VNM utility u, this yields the objective (2.2) with the risk-aversion coefficient $\alpha = -\frac{u''(\bar{v})}{u'(\bar{v})}$, which is constant over time.[4]

In Section 6 we also consider an alternative specification where, instead of mean-variance hedgers, hedgers live forever and maximize the utility function

$$-\mathrm{E}_t\left(\int_t^\infty e^{-\alpha c_{hs}} e^{-\rho_h(s-t)} ds\right) \tag{2.3}$$

where c_{hs} is consumption at $s \geq t$ and ρ_h is a subjective discount rate. We show in that section that our main results remain the same under that variation, but the resulting structure is much less tractable.

---

[4]The assumption that hedgers maximize a mean-variance objective over instantaneous changes in wealth simplifies our analysis and makes it possible to derive closed-form solutions. Our main results would remain the same under the alternative assumption that hedgers maximize a constant absolute risk aversion (CARA) utility over intertemporal consumption. We sketch this extension in Section 6.

Arbitrageurs maximize expected utility of intertemporal consumption. We assume time-additive utility and a constant coefficient of relative risk aversion (CRRA) $\gamma \geq 0$. When $\gamma \neq 1$, the arbitrageurs' objective at time t is

$$\mathrm{E}_t \left(\int_t^\infty \frac{c_s^{1-\gamma}}{1-\gamma} e^{-\rho(s-t)} ds \right), \tag{2.4}$$

where c_s is consumption at $s \geq t$ and ρ is a subjective discount rate. When $\gamma = 1$, the objective becomes

$$\mathrm{E}_t \left(\int_t^\infty \log(c_s) e^{-\rho(s-t)} ds \right). \tag{2.5}$$

Implicit in the definition of the arbitrageurs' objective for $\gamma > 0$, is that consumption is non-negative. The objective for $\gamma = 0$ can be defined for negative consumption, but we impose non-negativity as a constraint. Since negative consumption can be interpreted as a costly activity that arbitrageurs undertake to repay a loan, the non-negativity constraint can be interpreted as a collateral constraint: arbitrageurs cannot commit to engage in the costly activity, and can hence walk away from a loan not backed by collateral.

Our model can be given multiple interpretations. For example, it could represent the market for insurance against aggregate risks, e.g., weather or earthquakes. Under this interpretation, assets are insurance contracts and arbitrageurs are the insurers. Alternatively, the model could represent futures markets for commodities or financial assets, with arbitrageurs being the speculators. The model could also be interpreted more indirectly to represent stocks or bonds.

We present the equilibrium in steps. In the next Section we solve the model with only short-lived assets, i.e. claims on the next instant cash-flow. This variant results in simple and intuitive derivation of optimal risk-sharing, the price-of-risk and the wealth dynamics of hedgers and arbitrageurs, all which remain valid with standard, long-lived assets. We show this equivalence in Section 4 where we also present the expressions for expected return and return correlation for long-lived assets. In both of these Sections we work with the simplifying assumption that the risky assets are in zero net supply. That is, hedgers and arbitrageurs use risky assets only to share the aggregate risk represented by hedgers' endowment u. We reintroduce, arbitrary supply vector, s, in Section 6 and argue that our major results remain intact.

7

3 Short-lived assets

In this part, we consider the variant of the model where the only risky assets are short-lived. That is there are N risky assets in zero net supply which are contingent claims with infinitesimal maturity. Establishing a position z_t in the assets at time t costs $z_t^\top \pi_t dt$, and pays off $z_t^\top dD_t$ at time $t + dt$, where z_t and π_t are $N \times 1$ vectors. Since the assets are correlated with the hedgers' risky endowments, they can be traded to share risk. We define the return of the risky assets between t and $t + dt$ by $dR_t \equiv dD_t - \pi_t dt$. Eq. (2.1) implies that the instantaneous expected return is

$$\frac{\mathrm{E}_t(dR_t)}{dt} = \bar{D} - \pi_t, \tag{3.1}$$

and the instantaneous covariance matrix is

$$\frac{\mathrm{Var}_t(dR_t)}{dt} = \frac{\mathrm{E}_t(dR_t dR_t^\top)}{dt} = \sigma^\top \sigma = \Sigma. \tag{3.2}$$

Note that dR_t is also a return in excess of the riskless asset since investing $\pi_t dt$ in the riskless asset yields return $r\pi_t(dt)^2$, which is negligible relative to dR_t.

3.1 Equilibrium

We first solve the hedgers' maximization problem. Consider a hedger who holds a position x_t in the risky assets at time t. The change in the hedger's wealth between t and $t + dt$ is

$$dv_t = rv_t dt + x_t^\top (dD_t - \pi_t dt) + u^\top dD_t. \tag{3.3}$$

The first term in the right-hand side of (3.3) is the return from investing in the riskless asset, the second term is the return from investing in the risky assets, and the third term is the endowment. Substituting dD_t from (2.1) into (3.3), and the result into (2.2), we find the hedger's optimal asset demand.

Proposition 3.1 *The optimal policy of a hedger at time t is to hold a position*

$$x_t = \frac{\Sigma^{-1}(\bar{D} - \pi_t)}{\alpha} - u \tag{3.4}$$

in the risky assets.

8

The hedger's optimal demand for the risky assets consists of two components, which correspond to the two terms in the right-hand side of (3.4). The first term is the demand in the absence of the hedging motive. This demand consists of an investment in the tangent portfolio, scaled by the hedger's risk aversion coefficient α. The tangent portfolio is the inverse of the instantaneous covariance matrix Σ of asset returns times the vector $\bar{D} - \pi_t$ of instantaneous expected returns. The second term is the demand generated by the hedging motive. This demand consists of a short position in the portfolio u, which characterizes the sensitivity of hedgers' endowment to asset returns. Selling short an asset n for which u_n is positive hedges endowment risk.

We next study the arbitrageurs' maximization problem. Consider an arbitrageur who has wealth w_t at time t and holds a position y_t in the risky assets. The arbitrageur's budget constraint is

$$dw_t = rw_t dt + y_t^\top (dD_t - \pi_t dt) - c_t dt. \tag{3.5}$$

The first term in the right-hand side of (3.5) is the return from investing in the riskless asset, the second term is the return from investing in the risky assets, and the third term is consumption. The arbitrageur's value function depends not only on his own wealth w_t, but also on the total wealth of all arbitrageurs since the latter affects asset prices π_t. In equilibrium own wealth and total wealth coincide because all arbitrageurs hold the same portfolio and are in measure one. For the purposes of optimization, however, we need to make the distinction. We reserve the notation w_t for total wealth and denote own wealth by $\hat{w}_t$. We likewise use (c_t, y_t) for total consumption and position in the assets, and denote own consumption and position by $(\hat{c}_t, \hat{y}_t)$. We conjecture that the arbitrageur's value function is

$$V(\hat{w}_t, w_t) = q(w_t) \frac{\hat{w}_t^{1-\gamma}}{1-\gamma} \tag{3.6}$$

for $\gamma \neq 1$, and

$$V(\hat{w}_t, w_t) = \frac{1}{\rho} \log(\hat{w}_t) + q_1(w_t) \tag{3.7}$$

for $\gamma = 1$, where $q(w_t)$ and $q_1(w_t)$ are scalar functions of w_t. We set $q(w_t) = \frac{1}{\rho}$ for $\gamma = 1$. Substituting dD_t from (2.1) into (3.5), and the result into the arbitrageur's Bellman equation, we find the arbitrageur's optimal consumption and asset demand.

Proposition 3.2 *Given the value function (3.6) and (3.7), the optimal policy of an arbitrageur at time t is to consume*

$$\hat{c}_t = q(w_t)^{-\frac{1}{\gamma}} \hat{w}_t \tag{3.8}$$

and hold a position

$$\hat{y}_t = \frac{\hat{w}_t}{\gamma} \left(\Sigma^{-1}(\bar{D} - \pi_t) + \frac{q'(w_t) y_t}{q(w_t)} \right) \tag{3.9}$$

in the risky assets.

The arbitrageur's optimal consumption is proportional to his wealth $\hat{w}_t$, with the proportionality coefficient $q(w_t)^{-\frac{1}{\gamma}}$ being a function of total arbitrageur wealth w_t. The arbitrageur's optimal demand for the risky assets consists of two components, as for the hedgers. The first component is the demand in the absence of a hedging motive, and consists of an investment in the tangent portfolio, scaled by the arbitrageur's coefficient of absolute risk aversion $\frac{\gamma}{\hat{w}_t}$. The second component is the demand generated by intertemporal hedging (Merton (1973)). The arbitrageur hedges against changes in his investment opportunity set, and does so by holding a portfolio with weights proportional to the sensitivity of that set to asset returns. In our model the investment opportunity set is fully characterized by total arbitrageur wealth, and the sensitivity of that variable to asset returns is the average portfolio y_t of all arbitrageurs. Hence, the arbitrageur's hedging demand is a scaled version of y_t, as the second term in the right-hand side of (3.9) shows.

Since in equilibrium all arbitrageurs hold the same portfolio, both components of asset demand consist of an investment in the tangent portfolio. Setting $\hat{y}_t = y_t$ and $\hat{w}_t = w_t$ in (3.9), we find that the total asset demand of arbitrageurs is

$$y_t = \frac{\Sigma^{-1}(\bar{D} - \pi_t)}{A(w_t)}, \tag{3.10}$$

where

$$A(w_t) \equiv \frac{\gamma}{w_t} - \frac{q'(w_t)}{q(w_t)}. \tag{3.11}$$

Arbitrageurs' investment in the tangent portfolio is thus scaled by the coefficient $A(w_t)$, which measures effective risk aversion. Effective risk aversion is the sum of the static coefficient of absolute risk aversion $\frac{\gamma}{w_t}$, and of the term $-\frac{q'(w_t)}{q(w_t)}$ which corresponds to the intertemporal hedging demand.

10

Substituting the asset demand (3.4) of the hedgers and (3.10) of the arbitrageurs into the market-clearing equation

$$x_t + y_t = 0, \tag{3.12}$$

we find that asset prices π_t are

$$\pi_t = \bar{D} - \frac{\alpha A(w_t)}{\alpha + A(w_t)} \Sigma u. \tag{3.13}$$

Substituting (3.13) back into (3.10), we find that the arbitrageurs' position in the risky assets in equilibrium is

$$y_t = \frac{\alpha}{\alpha + A(w_t)} u. \tag{3.14}$$

Intuitively, hedgers want to sell the portfolio u to hedge their endowment. Arbitrageurs buy a fraction of that portfolio, and the rest remains with the hedgers. The fraction bought by arbitrageurs decreases in their effective risk aversion $A(w_t)$ and increases in the hedgers' risk aversion α, according to optimal risk-sharing. Expected asset returns are proportional to the covariance with the portfolio u, which is the single pricing factor in our model. The risk premium of that factor increases in the arbitrageurs' effective risk aversion, and is hence time-varying. The arbitrageurs' Sharpe ratio, defined as the expected return of their portfolio divided by the portfolio's standard deviation, also increases in their effective risk aversion. Using (3.13) and (3.14), we find that the Sharpe ratio is

$$SR_t \equiv \frac{y_t^\top (\bar{D} - \pi_t)}{\sqrt{y_t^\top \Sigma y_t}} = \frac{\alpha A(w_t)}{\alpha + A(w_t)} \sqrt{u^\top \Sigma u}. \tag{3.15}$$

Substituting the arbitrageurs' optimal policy from Proposition (3.2) into the Bellman equation, we can derive an ordinary differential equation (ODE) that the arbitrageurs' value function must satisfy.

Proposition 3.3 *If (3.6) is the value function for $\gamma \neq 1$, then $q(w_t)$ must solve the ODE*

$$\rho q = \gamma q^{1-\frac{1}{\gamma}} + \left(r - q^{-\frac{1}{\gamma}} \right) q' w + rq(1-\gamma) + \frac{1}{2} \left(q'' + \frac{2q'\gamma}{w} - \frac{2q'^2}{q} + \frac{q(1-\gamma)\gamma}{w^2} \right) \frac{\alpha^2}{\left(\alpha + \frac{\gamma}{w} - \frac{q'}{q} \right)^2} u^\top \Sigma u.$$

11

If (3.7) is the value function for $\gamma = 1$, then $q_1(w_t)$ must solve the ODE

$$\rho q_1 = \log(\rho) + \frac{r - \rho}{\rho} + (r - \rho)q_1' + \frac{1}{2}\left(q_1'' + \frac{2q_1'}{w} + \frac{1}{\rho w^2}\right)\frac{\alpha^2}{\left(\alpha + \frac{1}{w}\right)^2}u^\top \Sigma u.$$

(3.17)

3.2 Closed-Form Solutions

We next characterize the equilibrium more fully in two special cases: arbitrageurs have logarithmic preferences ($\gamma = 1$) and arbitrageurs are risk-neutral ($\gamma = 0$). A useful parameter in both cases is

$$z \equiv \frac{\alpha^2 u^\top \Sigma u}{2(\rho - r)}.$$

(3.18)

The parameter z is larger when hedgers are more risk averse (large α), or their endowment is riskier (large $u^\top \Sigma u$), or arbitrageurs are more patient (small ρ).

When $\gamma = 1$, (3.8) and $q(w_t) = \frac{1}{\rho}$ imply that arbitrageur consumption is equal to ρ times wealth. Eq. (3.11) implies that arbitrageur effective risk-aversion $A(w_t)$ is

$$A(w_t) = \frac{1}{w_t}.$$

(3.19)

Effective risk aversion is equal to the static coefficient of absolute risk aversion because the intertemporal hedging demand is zero.

When $\gamma = 0$, (3.8) implies that arbitrageur consumption is equal to zero in the region $q(w_t) > 1$ since $\frac{1}{\gamma} = \infty$. Moreover, $q(w_t) \geq 1$ since an arbitrageur can always consume his entire wealth $\hat{w}_t$ instantly and achieve utility $\hat{w}_t$. Therefore, there are two regions, one in which $q(w_t) > 1$ and arbitrageurs do not consume, and in which $q(w_t) = 1$ and arbitrageurs consume instantly until their total wealth w_t reaches the other region. The two regions are separated by a threshold $\bar{w} > 0$: for $w_t < \bar{w}$ arbitrageurs do not consume, and for $w_t > \bar{w}$ they consume instantly until w_t decreases to $\bar{w}$. The marginal utility $q(w_t)$ of an arbitrageur's wealth is high when the total wealth w_t of all arbitrageurs is low because liquidity provision is then more profitable. Arbitrageur effective risk-aversion $A(w_t)$ is the solution to a first-order ODE derived from (3.16). Proposition 3.20 solves this ODE in closed form in the limit when the riskless rate r goes to zero. For ease of exposition,

12

we refer from now on to the $r \to 0$ limit in the risk-neutral case as the "limit risk-neutral case." In subsequent sections we occasionally also take the $r \to 0$ limit in the logarithmic case, and refer to it as the "limit logarithmic case."

Proposition 3.4 *In the limit risk-neutral case ($\gamma = 0, r \to 0$), arbitrageur effective risk aversion is given by*

$$A(w_t) = \frac{\alpha}{1+z} \left(\sqrt{z} \cot \left(\frac{\alpha w_t}{\sqrt{z}} \right) - 1 \right) \tag{3.20}$$

for $w_t < \bar{w}$, and $A(w_t) = 0$ for $w_t \geq \bar{w}$, where the threshold $\bar{w}$ is given by

$$\cot \left(\frac{\alpha \bar{w}}{\sqrt{z}} \right) = \frac{1}{\sqrt{z}}. \tag{3.21}$$

The marginal utility of arbitrageur wealth is given by

$$q(w_t) = \exp \left\{ \frac{z}{1+z} \left[\log \sin \left(\frac{\alpha \bar{w}}{\sqrt{z}} \right) - \log \sin \left(\frac{\alpha w_t}{\sqrt{z}} \right) - \frac{\alpha}{1+z} (\bar{w} - w_t) \right] \right\} \tag{3.22}$$

for $w_t < \bar{w}$, and $q(w_t) = 1$ for $w_t \geq \bar{w}$.

Although arbitrageurs are risk-neutral, their effective risk aversion is positive in the region $w_t < \bar{w}$. This is because of the intertemporal hedging demand. Intuitively, arbitrageurs take into account that in states where their portfolio performs poorly, other arbitrageurs also perform poorly, and hence liquidity provision becomes more profitable. To have more wealth to provide liquidity in those states, arbitrageurs limit their investment in the risky assets, hence behaving as risk-averse.

Figure 1 plots arbitrageur effective risk aversion $A(w_t)$ as a function of wealth w_t. To choose values for α and $u^\top \Sigma u$, we set hedgers' initial wealth $\bar{v}$ to one: this is without loss of generality because we can redefine the numeraire. Since $\bar{v} = 1$, the parameter $\alpha = -\frac{u''(\bar{v})}{u'(\bar{v})}$ coincides with the hedgers' relative risk aversion coefficient, and we set it to 2. Moreover, the parameter $\sqrt{u^\top \Sigma u}$ coincides with the annualized standard deviation of the hedgers' endowment as a function of their initial wealth, and we set it to 15%. We set the arbitrageurs' subjective discount rate ρ to 4%, and the riskless rate r to 2%.

Figure 1 shows that in both the logarithmic and the risk-neutral cases, effective risk aversion $A(w_t)$ is decreasing and convex in arbitrageur wealth, and converges to infinity when wealth goes
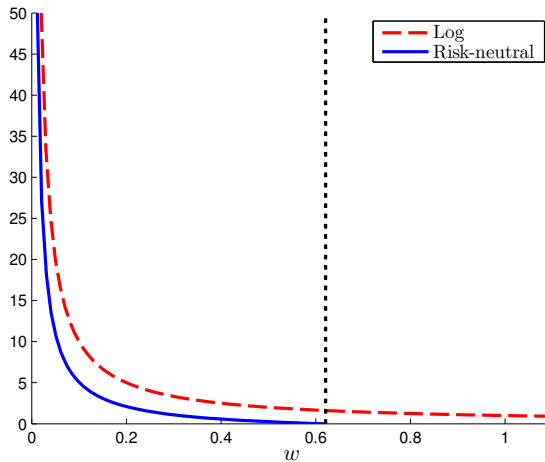
Figure 1: Arbitrageur effective risk aversion as a function of wealth in the logarithmic case (dashed line) and the risk-neutral case (solid line). Parameter values are $\alpha = 2$, $\sqrt{u^\top \Sigma u} = 15\%$, $\rho = 4\%$, and $r = 2\%$.

to zero. Moreover, effective risk aversion is smaller in the risk-neutral case than in the logarithmic case. These properties hold for all parameter values in the logarithmic case since $A(w_t) = \frac{1}{w_t}$. They also hold for all values of α, $u^\top \Sigma u$, and ρ in the limit risk-neutral case $(r \to 0)$, as we show in the proof of Proposition 3.4.

We next examine how changes in arbitrageur wealth affect expected asset returns and the arbitrageurs' positions and Sharpe ratio. When arbitrageurs are wealthier, they have lower effective risk aversion, and absorb a larger fraction of the portfolio u that hedgers want to sell. Arbitrageur positions are thus larger in absolute value: more positive for positive elements of u, which correspond to assets that hedgers want to sell, and more negative for negative elements of u, which correspond to assets that hedgers want to buy. Since arbitrageurs are less risk averse, they require smaller compensation for providing liquidity to hedgers. Expected asset returns, which measure this compensation, are thus smaller in absolute value. The same is true for the market prices of the Brownian risks, i.e., the expected returns per unit of risk exposure, and for the arbitrageurs' Sharpe ratio.

Proposition 3.5 *In both the logarithmic ($\gamma = 1$) and the limit risk-neutral ($\gamma = 0, r \to 0$) cases, an increase in arbitrageur wealth w_t:*

(i) *Raises the position of arbitrageurs in each asset in absolute value.*

(ii) *Lowers the expected return of each asset in absolute value.*

14

(iii) Lowers the market price of each Brownian risk in absolute value.

(iv) Lowers the arbitrageurs' Sharpe ratio.

We next derive the stationary distribution of arbitrageur wealth. Using this distribution, we can compute unconditional averages of endogenous variables, e.g., arbitrageurs' positions and Sharpe ratio.

Proposition 3.6 *If $z > 1$, then the stationary distribution of arbitrageur wealth has density*

$$d(w_t) = \frac{(\alpha w_t + 1)^2 w_t^{-\frac{1}{z}} \exp\left(-\frac{1}{2z}\left(\alpha^2 w_t^2 + 4\alpha w_t\right)\right)}{\int_0^\infty (\alpha w + 1)^2 w^{-\frac{1}{z}} \exp\left(-\frac{1}{2z}\left(\alpha^2 w^2 + 4\alpha w\right)\right) dw} \tag{3.23}$$

over the support $(0, \infty)$ in the logarithmic case ($\gamma = 1$), and density

$$d(w_t) = \frac{\left(\frac{\alpha + A(w_t)}{q(w_t)}\right)^2}{\int_0^{\bar{w}} \left(\frac{\alpha + A(w)}{q(w)}\right)^2 dw} \tag{3.24}$$

over the support $(0, \bar{w})$ in the limit risk-neutral case ($\gamma = 0, r \to 0$), where $A(w_t)$ and $q(w_t)$ are given by (3.20) and (3.22), respectively. If $0 < z < 1$, then wealth converges to zero in the long run, in both cases. If in the logarithmic case $z < 0$, then wealth converges to infinity in the long run.

The stationary distribution has a non-degenerate density if the parameter z defined by (3.18) is larger than one. This is the case when the hedgers' risk aversion α and endowment variance $u^\top \Sigma u$ are large, and the arbitrageurs' subjective discount rate ρ is small but exceeds the riskless rate r.

To provide an intuition for Proposition 3.6, we recall the standard Merton (1971) portfolio optimization problem in which an infinitely-lived investor with CRRA coefficient γ can invest in a riskless asset with instantaneous return r and in N risky assets with instantaneous expected excess return vector μ and covariance matrix Σ. The investor's wealth converges to infinity in the long run when

$$r + \frac{1}{2}\mu^\top \Sigma^{-1} \mu > \rho, \tag{3.25}$$

15

i.e., when the riskless rate plus one-half of the squared Sharpe ratio achieved from investing in the risky assets exceeds the investor's subjective discount rate ρ. When instead (3.25) holds in the opposite direction, wealth converges to zero. Intuitively, wealth converges to infinity when the investor accumulates wealth at a rate that exceeds sufficiently the rate at which he consumes.

Our model differs from the Merton problem because the arbitrageurs' Sharpe ratio is endogenously determined in equilibrium and decreases in their wealth (Proposition 3.5). Using (3.15) to substitute for the arbitrageurs' Sharpe ratio, we can write (3.25) as

$$ r + \frac{1}{2} \left(\frac{\alpha A(w_t)}{\alpha + A(w_t)} \right)^2 u^\top \Sigma u > \rho. \tag{3.26} $$

Transposing the result from the Merton problem thus suggests that there are three possibilities for the long-run dynamics. If (3.26) is satisfied for all values of w_t, then wealth converges to infinity. If (3.26) is violated for all values of w_t, then wealth converges to zero. If, finally, (3.26) is violated for large values but is satisfied for values close to zero, neither convergence occurs and wealth has a non-degenerate stationary density. Intuitively, a density can exist because the dynamics of arbitrageur wealth are self-correcting: when wealth becomes close to zero the Sharpe ratio increases and (3.26) becomes satisfied, and when wealth becomes large the Sharpe ratio decreases and (3.26) becomes violated.

When $\rho < r$, and so $z < 0$, (3.26) is satisfied for all values of w_t. Therefore, w_t converges to infinity. When $\rho > r$, and so $z > 0$, (3.26) is violated for values of w_t close to its upper bound (infinity in the logarithmic case and $\bar{w}$ in the risk-neutral case) because $A(w_t)$ is close to zero for those values. Therefore, w_t either converges to zero or has a non-degenerate stationary density. Convergence to zero occurs if (3.26) is violated for w_t close to zero because it is then violated for all values of w_t. Since $A(w_t)$ is close to infinity for w_t close to zero, w_t converges to zero exactly when $z < 1$. Intuitively, wealth converges to zero when α and $u^\top \Sigma u$ are small because then arbitrageurs earn low expected returns for providing liquidity to hedgers. When instead $z > 1$, w_t has a non-degenerate stationary density. Proposition 3.7 characterizes the shape of that density.

Proposition 3.7 *Suppose that $z > 1$. The density $d(w_t)$ of the stationary distribution:*

(i) Is decreasing in w_t if $z < \frac{27}{8}$ in the logarithmic case $(\gamma = 1)$ and if $z < 4$ in the limit risk-neutral case $(\gamma = 0, r \to 0)$.

(ii) Is bimodal in w_t otherwise. That is, it is decreasing in w_t for $0 < w_t < \bar{w}_1$, increasing in

16

w_t for $\bar{w}_1 < w_t < \bar{w}_2$, and again decreasing in w_t for $w_t > \bar{w}_2$. In the logarithmic case, the thresholds $\bar{w}_1 < \bar{w}_2$ are the two positive roots of

$$(\alpha w)^3 + 3(\alpha w)^2 + (3 - 2z)\alpha w + 1 = 0. \tag{3.27}$$

In the limit risk-neutral case, they are given by

$$A(\bar{w}_1) \equiv \alpha \frac{z - 2 + \sqrt{z(z-4)}}{2}, \tag{3.28}$$

$$A(\bar{w}_2) \equiv \alpha \frac{z - 2 - \sqrt{z(z-4)}}{2}, \tag{3.29}$$

where $A(w_t)$ is given by (3.20), and they satisfy $0 < \bar{w}_1 < \bar{w}_2 < \bar{w}$.

(iii) Shifts to the right in the monotone likelihood ratio sense when α or $u^\top \Sigma u$ increase, in both the logarithmic and the limit risk-neutral cases.

The shape of the stationary density is fully determined by the parameter z. When z is not much larger than one, the density is decreasing, and so values close to zero are more likely than larger values. When instead z is sufficiently larger than one, the density becomes bimodal, with the two maxima being zero and an interior point $\bar{w}_2$ of the support. Values close to these maxima are more likely than intermediate values, meaning that the system spends more time at these values than in the middle. The intuition is that when the hedgers' risk aversion α and endowment variance $u^\top \Sigma u$ are large, arbitrageurs earn high expected returns for providing liquidity, and their wealth grows fast. Therefore, large values of w_t are more likely in steady state than intermediate values. At the same time, while expected returns are highest when wealth is small, wealth grows away from small values slowly in absolute terms. Therefore, small values of w_t are more likely than intermediate values.

Figure 2 plots the stationary density in the logarithmic and risk-neutral cases. The solid lines are drawn for the same parameter values as in Figure 1. The dashed lines are drawn for the same values except that hedger risk aversion α is raised from 2 to 4. The solid lines are decreasing in wealth, while the dashed lines are bimodal. These patterns are consistent with Proposition 3.7 since z is equal to 2.25 for the solid lines and to 9 for the dashed lines.

We next perform comparative statics with respect to the hedgers' risk aversion α and endowment variance $u^\top \Sigma u$. We perform "conditional" comparative statics, where we compute how changes in α and $u^\top \Sigma u$ affect endogenous variables, conditionally on a given level of arbitrageur
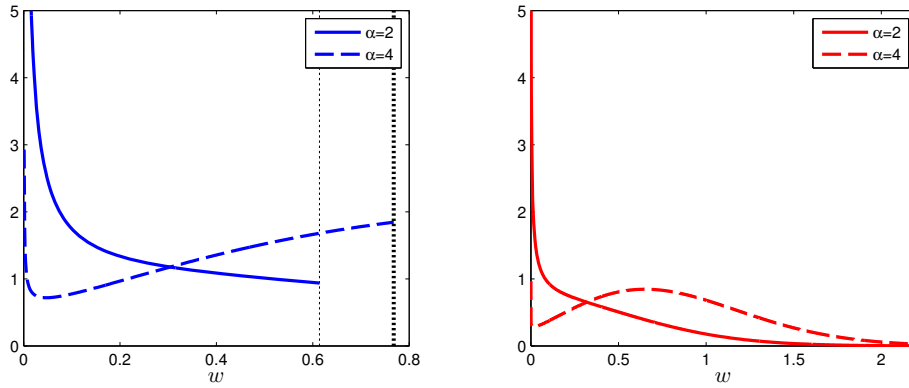
17

Figure 2: Stationary density of arbitrageur wealth in the logarithmic case (right panel) and the risk-neutral case (left panel). The solid lines are drawn for $\alpha = 2$, $\sqrt{u^\top \Sigma u} = 15\%$, $\rho = 4\%$, and $r = 2\%$. The dashed lines are drawn for the same values except that $\alpha = 4$.

wealth. We also perform "unconditional" comparative statics, where we compute how changes in α and $u^\top \Sigma u$ affect unconditional averages of the endogenous variables under the stationary distribution of wealth. The two types of comparative statics differ sharply.

Proposition 3.8 *Conditionally on a given level w_t of arbitrageur wealth, the following comparative statics hold:*

(i) *An increase in the hedgers' risk aversion α raises the arbitrageurs' Sharpe ratio. In the logarithmic case ($\gamma = 1$), the position of arbitrageurs in each asset increases in absolute value. In the limit risk-neutral case ($\gamma = 0, r \to 0$), the position of arbitrageurs in each asset decreases in absolute value, except when w_t is below a threshold, which is negative if $z < 1$.*

(ii) *An increase in the variance $u^\top \Sigma u$ of hedgers' endowment raises the arbitrageurs' Sharpe ratio. In the logarithmic case, arbitrageur positions do not change. In the limit risk-neutral case, the position of arbitrageurs in each asset decreases in absolute value.*

Result (i) of Proposition 3.8 concerns changes in hedger risk aversion. One would expect that when hedgers become more risk averse, they transfer more risk to arbitrageurs. This result holds in the logarithmic case, but the opposite result can hold in the risk-neutral case. This is because an increase in hedger risk aversion can generate an even larger increase in arbitrageur effective risk aversion through an increase in the intertemporal hedging demand. Recall that risk-neutral arbitrageurs behave as risk-averse because they seek to preserve wealth in states where other

18

arbitrageurs realize losses and liquidity provision becomes more profitable. When hedgers are more risk averse, this effect becomes stronger because liquidity provision becomes more profitable for each level of arbitrageur wealth and more sensitive to changes in wealth. The effect is not present in the logarithmic case because effective risk aversion is equal to the static coefficient of absolute risk aversion, which depends only on wealth. In both the logarithmic and the risk-neutral cases, an increase in hedger risk aversion raises the Sharpe ratio of arbitrageurs because the expected return on their portfolio increases.

Result (ii) of Proposition 3.8 concerns changes in the variance of hedgers' endowment. In the logarithmic case, such changes do not affect arbitrageur effective risk aversion and positions. In the risk-neutral case, however, there is an effect, which parallels that of hedger risk aversion. When the variance is high, e.g., because asset cashflows dD_t are more volatile, liquidity provision becomes more profitable. As a consequence, arbitrageurs have higher effective risk aversion and hold smaller positions. In both the logarithmic and the risk-neutral cases, an increase in variance raises the arbitrageurs' Sharpe ratio.

The results of Proposition 3.8 can be related to findings of recent empirical papers that measure the demand or supply of different investor groups and its relationship with expected returns. Examples are Hong and Yogo (2012) and Chen, Joslin, and Ni (2013) for derivatives markets, and Buyuksahin and Robe (2010), Irwin and Sanders (2010), Tang and Xiong (2010), Hamilton and Wu (2011) and Cheng, Kirilenko, and Xiong (2012) for commodity markets. These papers typically interpret shocks that reduce investor positions and increase expected returns as downward shifts to supply that could result from tighter constraints of liquidity providers. Intuitively, when providers become more constrained, they take smaller positions and require larger expected returns as compensation. These effects can be derived within our model if we identify tighter constraints with a reduction in arbitrageur wealth (Proposition 3.5). Proposition 3.8 suggests, however, that the same effects can arise following upward shifts to demand. For example, Result (i) shows that in the risk-neutral case an increase in hedgers' risk aversion can lower arbitrageurs' positions and raise expected returns.

We next turn to unconditional comparative statics. Figure 3 plots the unconditional Sharpe ratio of arbitrageurs as a function of α (left panel) and $u^\top \Sigma u$ (right panel). The results are in sharp contrast to the conditional comparative statics. While an increase in α and $u^\top \Sigma u$ raises the Sharpe ratio conditionally on a given level of wealth (Proposition 3.8), it can lower it when comparing unconditional averages. Intuitively, for larger values of α and $u^\top \Sigma u$, arbitrageur wealth

19

grows faster, and its stationary density shifts to the right (Proposition 3.7). Therefore, while the conditional Sharpe ratio increases, the unconditional one can decrease because high values of wealth, which yield low Sharpe ratios, become more likely.
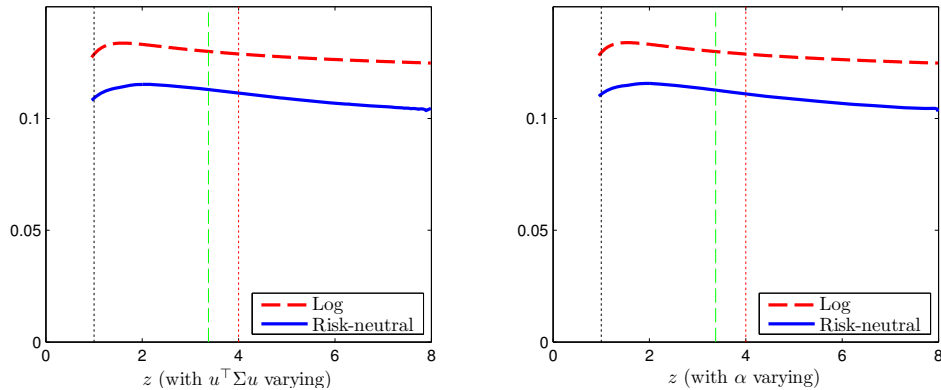

Figure 3: The unconditional Sharpe ratio of arbitrageurs as a function of α (left panel) and $u^\top \Sigma u$ (right panel), for the logarithmic case (dashed lines) and the risk-neutral case (solid lines). When α varies, the remaining parameters are set to $\sqrt{u^\top \Sigma u} = 15\%$, $\rho = 4\%$, and $r = 2\%$. When $u^\top \Sigma u$ varies, the remaining parameters are set to $\alpha = 2$, $\rho = 4\%$, and $r = 2\%$. The left-most vertical bar is the threshold $z = 1$ beyond which the stationary distribution has a non-degenerate density. The vertical bars to the right are the thresholds $z = \frac{27}{8}$ and $z = 4$ beyond which the density becomes bimodal in the logarithmic and in the limit risk-neutral cases, respectively.

4 Long-Lived Assets

In this section we replace the short-lived assets by N long-lived assets that pay the infinite stream of the short-lived assets' cashflows. We maintain the assumption of zero net supply, but sketch how our analysis can be extended to positive supply in Section 6. We show that risk-sharing and the dynamics of arbitrageur wealth remain the same as with short-lived assets. With long-lived assets, however, we can study a richer set of issues, which include the time-variation in volatilities and correlations, and the measurement and pricing of liquidity risk. These issues arise with long-lived assets because their return is sensitive to changes in arbitrageur wealth and liquidity is related to wealth.

Establishing a position Z_t in the long-lived assets at time t costs $Z_t^\top S_t dt$, and pays off $Z_t^\top (dD_t + S_{t+dt})$ at time $t + dt$, where Z_t and S_t are $N \times 1$ vectors. We conjecture that in equilibrium S_t follows the Ito process

$$dS_t = \mu_{St} dt + \sigma_{St}^\top dB_t, \tag{4.1}$$

where μ_{St} is a $N \times 1$ vector and σ_{St} is a $N \times N$ matrix.

We define the return of the long-lived assets between t and $t + dt$, in excess of the riskless asset, by $dR_t \equiv dS_t + dD_t - rS_t dt$. Eqs. (2.1) and (4.1) imply that the instantaneous expected return is

$$\frac{\mathrm{E}_t(dR_t)}{dt} = \mu_{St} + \bar{D} - rS_t, \tag{4.2}$$

and the instantaneous covariance matrix is

$$\frac{\mathrm{Var}_t(dR_t)}{dt} = (\sigma_{St} + \sigma)^\top (\sigma_{St} + \sigma). \tag{4.3}$$

4.1 Equilibrium

Eq. (3.3), which characterizes the change in a hedger's wealth between t and $t + dt$, is replaced by

$$dv_t = rv_t dt + X_t^\top (dS_t + dD_t - rS_t dt) + u^\top dD_t, \tag{4.4}$$

where X_t denotes the hedger's position in the long-lived assets. Eq. (3.5), which characterizes the change in an arbitrageur's wealth over the same interval, is similarly replaced by

$$dw_t = (rw_t - c_t)dt + Y_t^\top (dS_t + dD_t - rS_t dt), \tag{4.5}$$

where Y_t denotes the arbitrageur's position. Because the market is complete under long-lived assets, as it is under short-lived assets, the two asset structures generate the same allocation of risk.

Lemma 4.1 *An equilibrium (S_t, X_t, Y_t) with long-lived assets can be constructed from an equilibrium (π_t, x_t, y_t) with short-lived assets by:*

(i) Choosing the price process S_t such that

$$\left(\sigma^\top\right)^{-1} (\bar{D} - \pi_t) = \left((\sigma_{St} + \sigma)^\top\right)^{-1} (\mu_{St} + \bar{D} - rS_t). \tag{4.6}$$

(ii) Choosing the asset positions X_t of hedgers and Y_t of arbitrageurs such that

$$\sigma x_t = (\sigma_{St} + \sigma)X_t, \tag{4.7}$$

$$\sigma y_t = (\sigma_{St} + \sigma)Y_t. \tag{4.8}$$

21

In the equilibrium with long-lived assets the dynamics of arbitrageur wealth, the arbitrageurs' Sharpe ratio, and the exposures of hedgers and arbitrageurs to the Brownian shocks, are the same as in the equilibrium with short-lived assets.

Eqs. (4.7) and (4.8) construct positions of hedgers and arbitrageurs in the long-lived assets so that the exposures to the underlying Brownian shocks are the same as with short-lived assets. Eq. (4.6) constructs a price process such that the market prices of the Brownian risks are also the same. Given this price process, agents choose optimally the risk exposures in (4.7) and (4.8), and markets clear.

The price S_t is a function of arbitrageur wealth w_t only. Using Ito's lemma to compute the drift μ_{St} and diffusion σ_{St} of the price process as a function of the dynamics of w_t, and substituting into (4.6), we can determine $S(w_t)$ up to an ODE.

Proposition 4.1 *The price of the long-lived assets is given by*

$$S(w_t) = \frac{\bar{D} - \alpha \Sigma u}{r} + g(w_t)\Sigma u, \tag{4.9}$$

where the scalar function $g(w_t)$ satisfies the ODE

$$\left(r - q^{-\frac{1}{\gamma}}\right)wg' + \frac{\alpha^2}{2(\alpha + A)^2}u^\top \Sigma u g'' - rg = -\frac{\alpha^2}{\alpha + A}. \tag{4.10}$$

The price in (4.9) is the sum of two terms. The first term, $\frac{\bar{D} - \alpha \Sigma u}{r}$, is the price that would prevail in the absence of arbitrageurs. Indeed, if hedgers were the only traders in a market with short-lived assets, their demand (3.4) would equal the asset supply, which is zero. Solving for the market-clearing price yields $\pi_t = \bar{D} - \alpha \Sigma u$. Long-lived assets would trade at the present value of the infinite stream of these prices discounted at the riskless rate r, which is $\frac{\bar{D} - \alpha \Sigma u}{r}$. The second term, $g(w_t)\Sigma u$, measures the price impact of arbitrageurs. Since arbitrageurs buy a fraction of the portfolio u that hedgers want to sell, they cause assets covarying positively with that portfolio to become more expensive. Therefore, the function $g(w_t)$ should be positive, and equal to zero for $w_t = 0$. Moreover, since arbitrageurs have a larger impact the wealthier they are, $g(w_t)$ should be increasing in w_t, as we confirm in the special cases studied in Section 4.2.

Expected asset returns and the covariance matrix of returns are driven by the sensitivity of the price to changes in arbitrageur wealth w_t. Therefore, they are driven by the term $g(w_t)\Sigma u$ and do

22

not depend on $\frac{\bar{D}-\alpha\Sigma u}{r}$. In the proof of Proposition 4.1 we show that the instantaneous expected return is

$$\frac{\mathrm{E}_t(dR_t)}{dt} = \frac{\alpha A(w_t)}{\alpha + A(w_t)} \left[f(w_t)u^\top\Sigma u + 1 \right] \Sigma u, \tag{4.11}$$

and the instantaneous covariance matrix is

$$\frac{\mathrm{Var}_t(dR_t)}{dt} = f(w_t) \left[f(w_t)u^\top\Sigma u + 2 \right] \Sigma uu^\top\Sigma + \Sigma, \tag{4.12}$$

where

$$f(w_t) \equiv \frac{\alpha g'(w_t)}{\alpha + A(w_t)}. \tag{4.13}$$

The covariance matrix (4.12) is the sum of a "fundamental" component Σ, driven purely by shocks to assets' underlying cashflows dD_t, and an "endogenous" component $f(w_t) \left[f(w_t)u^\top\Sigma u + 2 \right] \Sigma uu^\top\Sigma$, introduced because cashflow shocks affect arbitrageur wealth w_t which affects prices. Endogenous risk is zero in the case of short-lived assets because their payoff dD_t is not sensitive to changes in w_t. Changes in w_t, however, affect the payoff $dD_t + S_{t+dt}$ of long-lived assets because they affect the price S_{t+dt}. Therefore, endogenous risk arises with long-lived assets, and we show that it drives the patterns of volatilities, correlations, and expected returns.

The effect of w_t on prices is proportional to the covariance Σu with the portfolio u. Therefore, the endogenous covariance between assets n and n' is proportional to the product between the elements n and n' of the vector Σu. Expected returns are proportional to Σu, as in the case of short-lived assets. The proportionality coefficient is different than in that case, however, because it is influenced by the endogenous covariance.

4.2 Closed-Form Solutions

We next characterize the equilibrium more fully in the logarithmic case ($\gamma = 1$) and the risk-neutral case ($\gamma = 0$). We compute the function $g'(w_t)$ that characterizes the sensitivity of the price to changes in arbitrageur wealth in closed form in the limit when the riskless rate r goes to zero. In that limit the price is not well defined since the constant term $\frac{\bar{D}-\alpha\Sigma u}{r}$ converges to infinity. The function $g(w_t)$ is well-defined, however, and so are expected asset returns and the covariance matrix

of returns. Hence, as long as $g(w_t)$ is continuous with respect to r, our results are informative about the properties of these quantities close to the limit, where the price is well defined.

Proposition 4.2 *The function $g'(w_t)$ is given by*

$$g'(w_t) = \frac{2w_t^{\frac{1}{z}} \exp\left(\frac{1}{2z}\left(\alpha^2 w_t^2 + 4\alpha w_t\right)\right)}{u^\top \Sigma u} \int_{w_t}^\infty \left(\alpha + \frac{1}{w}\right) w^{-\frac{1}{z}} \exp\left(-\frac{1}{2z}\left(\alpha^2 w^2 + 4\alpha w\right)\right) dw \quad (4.14)$$

for $w_t \in (0, \infty)$ in the limit logarithmic case ($\gamma = 1, r \to 0$), and by

$$g'(w_t) = \frac{2z}{(1+z)u^\top \Sigma u}\left[\log\sin\left(\frac{\alpha \bar{w}}{\sqrt{z}}\right) - \log\sin\left(\frac{\alpha w_t}{\sqrt{z}}\right) + \alpha(\bar{w} - w_t)\right] \quad (4.15)$$

for $w_t \in (0, \bar{w})$ in the limit risk-neutral case ($\gamma = 0, r \to 0$). In both cases $g'(w_t) > 0$.

We next examine how changes in arbitrageur wealth affect expected asset returns, volatilities, correlations, and arbitrageur positions.

Proposition 4.3 *An increase in arbitrageur wealth w_t has the following effects in both the limit logarithmic ($\gamma = 1, r \to 0$) and the limit risk-neutral ($\gamma = 0, r \to 0$) cases:*

(i) *A hump-shaped effect on the expected return of each asset, in absolute value, except when $z < \frac{1}{2}$ in the logarithmic case, where the effect is decreasing. The hump peaks at a value $\bar{w}_a$ that is common to all assets.*

(ii) *A hump-shaped effect on the volatility of the return of each asset. The hump peaks at a value $\bar{w}_b$ that is common to all assets and is larger than the corresponding value $\bar{w}_a$ for expected return.*

(iii) *The same hump-shaped effect as in Part (ii) on the covariance between the returns of each asset pair (n, n') if $(\Sigma u)_n(\Sigma u)_{n'} > 0$, and the opposite, i.e., inverse hump-shaped effect, if $(\Sigma u)_n(\Sigma u)_{n'} < 0$.*

(iv) *The same hump-shaped effect as in Part (ii) on the correlation between the returns of each asset pair (n, n') if*

$$\frac{(\Sigma u)_n(\Sigma u)_{n'}\Sigma_{nn} - (\Sigma u)_n^2 \Sigma_{nn'}}{f(w_t)\left[f(w_t)u^\top \Sigma u + 2\right](\Sigma u)_n^2 + \Sigma_{nn}} + \frac{(\Sigma u)_n(\Sigma u)_{n'}\Sigma_{n'n'} - (\Sigma u)_{n'}^2 \Sigma_{nn'}}{f(w_t)\left[f(w_t)u^\top \Sigma u + 2\right](\Sigma u)_{n'}^2 + \Sigma_{n'n'}} > 0, \quad (4.16)$$

and the opposite, i.e., inverse hump-shaped, effect if (4.16) holds in the opposite direction.

24

(v) An increasing effect on the position of arbitrageurs in each asset, in absolute value.

Since the fundamental component Σ of the covariance matrix is independent of arbitrageur wealth, the hump-shaped patterns of volatilities, covariances, and correlations are driven by the endogenous component. The intuition for the hump shape in the case of volatilities can be seen by computing the diffusion of the price process. Ito's lemma implies that $\sigma_{St} = \sigma_{wt}S'(w_t)^\top$, i.e., price volatility (diffusion) is equal to the volatility of arbitrageur wealth times the sensitivity of the price to changes in wealth. The volatility of wealth is increasing in wealth, and converges to zero when wealth goes to zero. Intuitively, when arbitrageurs are poor, they hold small positions and take almost no risk. The sensitivity of price to changes in wealth is instead decreasing in wealth, and converges to zero when wealth becomes large (close to infinity in the logarithmic case and to $\bar{w}$ in the risk-neutral case). Intuitively, when arbitrageurs are wealthy, they provide perfect liquidity to hedgers, and changes to their wealth have no impact on price. Therefore, price volatility converges to zero at both extremes of the wealth distribution, and this accounts for the hump-shaped pattern of return volatilities.[5]

The intuition for the hump shape in the case of covariances is similar to that for volatilities. Price movements caused by changes in arbitrageur wealth are small at the extremes of the wealth distribution and larger in the middle. This yields a hump-shaped pattern for the covariance between two assets n and n', if the prices of these assets move in the same direction. Movements are in the same direction when the term (n, n') of the endogenous covariance matrix is positive. This term is equal to $(\Sigma u)_n (\Sigma u)_{n'}$, and is likely to be positive when the corresponding components of the vector u have the same sign, i.e., arbitrageurs either buy both assets from the hedgers or sell both assets to them. When, for example, both assets are bought by arbitrageurs, they both appreciate when arbitrageur wealth go up, yielding a positive covariance.

The effect on correlations is more complicated than that on covariances because it includes the effect on volatilities. Suppose that changes in arbitrageur wealth move the prices of assets n and n' in the same direction, and hence have a hump-shaped effect on their covariance. Because, however, the effect on volatilities, which are in the denominator, is also hump-shaped, the overall effect on the correlation can be inverse hump-shaped. Intuitively, arbitrageurs can cause assets to become less correlated because the increase in volatilities that they cause can swamp the increase

---

[5]Price volatility converges to zero at the extremes of the wealth distribution because we are assuming for simplicity *i.i.d.* cashflows dD_t. Under *i.i.d.* cashflows, a cashflow shock does not have a direct effect on prices, i.e., does not affect prices holding arbitrageur wealth constant. Return volatility remains positive even at the extremes of the wealth distribution because a cashflow shock has a direct effect on returns. Under a persistent cashflow process, price volatility would not converge to zero at the extremes, while also remaining hump-shaped.

in covariance.

The hump-shaped pattern of expected returns derives from that of volatilities. Expected returns per unit of risk exposure, i.e., the market prices of the Brownian risks, are the same as in the equilibrium with short-term assets, and are hence decreasing in wealth (Proposition 3.5). But because the volatility of long-term assets is hump-shaped in wealth, their expected returns are generally also hump-shaped.

Figures 4 and 5 illustrate the behavior of assets' Sharpe ratios, expected returns, volatilities, and correlations as a function of arbitrageur wealth in the logarithmic and risk-neutral cases, respectively. The figures are drawn for the same parameter values as in Figure 2.
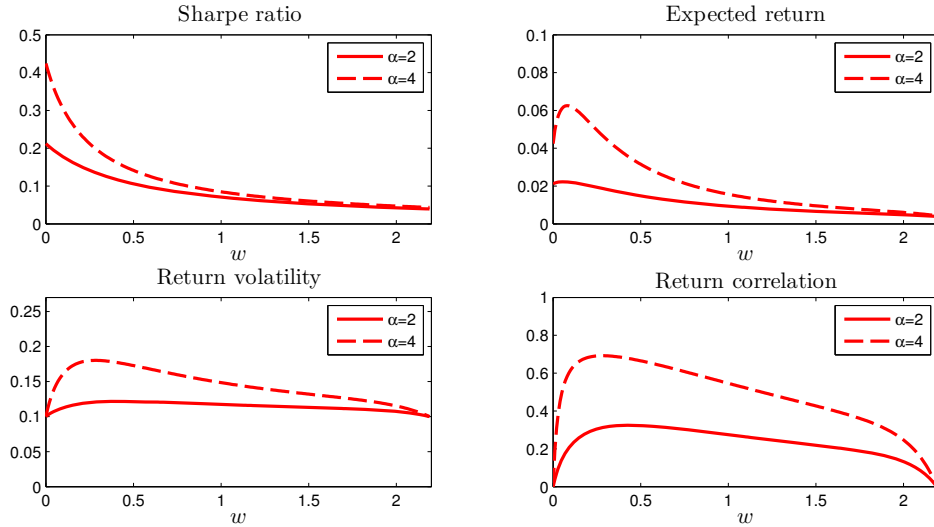

Figure 4: Assets' Sharpe ratios, expected returns, volatilities, and correlations as a function of arbitrageur wealth in the logarithmic case. The solid lines are drawn for $\alpha = 2$, $\sqrt{u^\top \Sigma u} = 15\%$, $\rho = 4\%$, $r = 2\%$, $N = 2$ symmetric assets with independent cashflows, and $\Sigma_{11} = \Sigma_{22} = 10\%$. The dashed lines are drawn for the same values except that $\alpha = 4$.

Using Figures 4 and 5, we can compare the logarithmic and risk-neutral cases. The assets' Sharpe ratios are higher in the logarithmic case, as one would expect since risk aversion is higher. Expected returns, however, can be higher in the risk-neutral case (as the figures show more clearly for $\alpha = 4$). This effect derives from volatilities, whose endogenous component can be higher in the risk-neutral case.
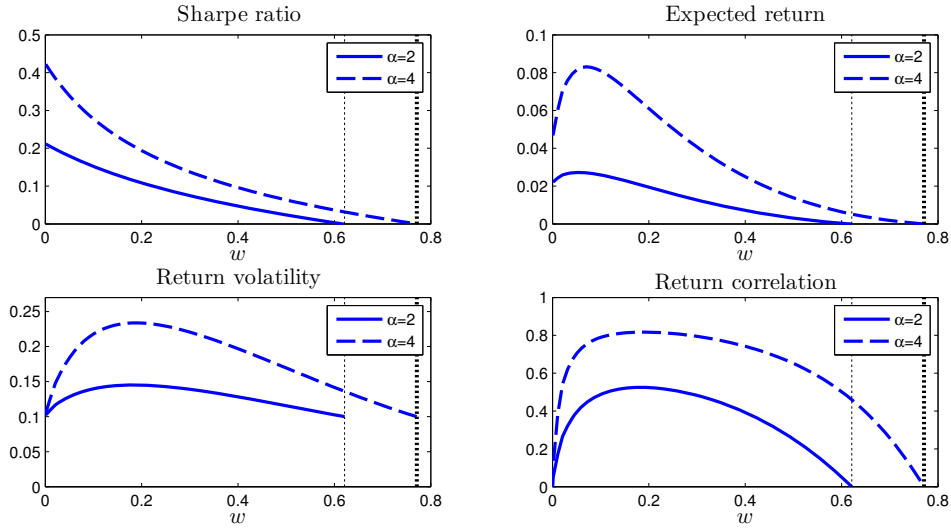
26

Figure 5: Assets' Sharpe ratios, expected returns, volatilities, and correlations as a function of arbitrageur wealth in the risk-neutral case. The solid lines are drawn for $\alpha = 2$, $\sqrt{u^\top \Sigma u} = 15\%$, $\rho = 4\%$, $r = 2\%$, $N = 2$ symmetric assets with independent cashflows, and $\Sigma_{11} = \Sigma_{22} = 10\%$. The dashed lines are drawn for the same values except that $\alpha = 4$.

5   Liquidity Risk

In this section we explore the implications of our model for liquidity risk. We assume long-lived assets, as in the previous section. We define liquidity based on the impact that hedgers have on prices. Consider an increase in the parameter u_n that characterizes hedgers' willingness to sell asset n. This triggers a decrease $\frac{\partial X_{nt}}{\partial u_n}$ in the quantity of the asset held by the hedgers, and a decrease $\frac{\partial S_{nt}}{\partial u_n}$ in the asset price. Asset n has low liquidity if the price change per unit of quantity traded is large. That is, the illiquidity of asset n is defined by

$$\lambda_{nt} \equiv \frac{r \frac{\partial S_{nt}}{\partial u_n}}{\frac{\partial X_{nt}}{\partial u_n}}, \tag{5.1}$$

where we multiply by r to ensure a well-behaved limit for our closed-form solutions. The measure (5.1) is in the spirit of Kyle (1985) and Amihud (2002).

A drawback of the measure (5.1) in the context of our model is that u_n is constant over time, and hence λ_{nt} cannot be computed by an empiricist. One interpretation of (5.1) is that there are small shocks to u_n, which an empiricist can observe and use to compute λ_{nt}. In Section 6 we sketch how our analysis can be extended to stochastic u and to additional measures of liquidity used

in empirical work. An alternative interpretation of (5.1) is cross-sectional, as a price differential between pairs of assets that are identical in cashflow variance and covariance with other assets, and differ only in their respective components of u.

Proposition 5.1 *Illiquidity λ_{nt} is equal to*

$$\left(1 + \frac{A(w_t)}{\alpha} + g'(w_t)u^\top \Sigma u\right)(\alpha - rg(w_t))\Sigma_{nn}. \tag{5.2}$$

An increase in arbitrageur wealth w_t lowers illiquidity in both the limit logarithmic $(\gamma = 1, r \to 0)$ and the limit risk-neutral $(\gamma = 0, r \to 0)$ cases.

Proposition 5.1 identifies a time-series and a cross-sectional dimension of illiquidity. In the time-series, illiquidity varies in response to changes in arbitrageur wealth, and is a decreasing function of wealth. This variation is common across assets and corresponds to the two terms in parentheses in (5.2). In the cross-section, illiquidity is higher for assets with more volatile cashflows. The dependence of illiquidity on the asset index n is through the asset's cashflow variance Σ_{nn}, the last term in (5.2). The time-series and cross-sectional dimensions of illiquidity interact: assets with more volatile cashflows have higher illiquidity for any given level of wealth, and the time-variation of their illiquidity is more pronounced.

Using Proposition 5.1, we can compute the covariance between asset returns and aggregate illiquidity. Since illiquidity varies over time because of arbitrageur wealth, and with an inverse relationship, the covariance of the return vector with illiquidity is equal to the covariance with wealth times a negative coefficient. Proposition 4.1 implies, in turn, that the covariance of the return vector with wealth is proportional to Σu. This is the covariance between asset cashflows and the cashflows of the portfolio u, which characterizes hedgers' supply. The intuition for the proportionality is that when arbitrageurs realize losses, they sell a fraction of u, and this lowers asset prices according to the covariance with u. Therefore, the covariance between asset returns and aggregate illiquidity $\Lambda_t \equiv \frac{\sum_{n=1}^{N} \lambda_{nt}}{N}$ is

$$\frac{\text{Cov}_t(d\Lambda_t, dR_t)}{dt} = C^\Lambda(w_t)\Sigma u, \tag{5.3}$$

where $C^\Lambda(w_t)$ is a negative coefficient. Assets that suffer the most when aggregate illiquidity increases and arbitrageurs sell a fraction of the portfolio u, are those corresponding to large components $(\Sigma u)_n$ of Σu. They have volatile cashflows (high Σ_{nn}), or are in high supply by hedgers (high u_n), or correlate highly with assets with those characteristics.

Using Proposition 5.1, we can compute two additional liquidity-related covariances: the covariance between an asset's illiquidity and aggregate illiquidity, and the covariance between an asset's illiquidity and aggregate return. We take the aggregate return to be that of the portfolio u, which characterizes hedgers' supply. Acharya and Pedersen (2005) show theoretically, within a model with exogenous transaction costs, that both covariances are linked to expected returns in the cross-section. In our model, the time-variation of an asset's illiquidity is proportional to the asset's cashflow variance Σ_{nn}. Therefore, the covariances between the asset's illiquidity on one hand, and aggregate illiquidity or return on the other, are proportional to Σ_{nn}.

Corollary 5.1 *In the cross-section of assets:*

(i) *The covariance between asset n's return dR_{nt} and aggregate illiquidity Λ_t is proportional to the covariance $(\Sigma u)_n$ between the asset's cashflows and the cashflows of the hedger-supplied portfolio u.*

(ii) *The covariance between asset n's illiquidity λ_{nt} and aggregate illiquidity Λ_t is proportional to the variance Σ_{nn} of the asset's cashflows.*

(iii) *The covariance between asset n's illiquidity λ_{nt} and aggregate return $u^\top dR_t$ is proportional to the variance Σ_{nn} of the asset's cashflows.*

The proportionality coefficients are negative, positive, and negative, respectively, in both the limit logarithmic ($\gamma = 1, r \to 0$) and the limit risk-neutral ($\gamma = 0, r \to 0$) cases.

We next determine the link between liquidity-related covariances and expected returns. Recall from (4.11) that the expected return of asset n is proportional to $(\Sigma u)_n$. This is exactly proportional to the covariance between the asset's return and aggregate illiquidity. Thus, aggregate illiquidity is a priced risk factor that explains expected returns perfectly. Intuitively, expected returns are priced from the portfolio of arbitrageurs, who are the marginal agents. Moreover, the covariance between asset returns and aggregate illiquidity identifies perfectly the arbitrageurs' portfolio. This is because (i) changes in aggregate illiquidity are driven by arbitrageur wealth, and (ii) the portfolio of trades that arbitrageurs perform when their wealth changes is proportional to their existing portfolio and impacts returns proportionately to the covariance with that portfolio.

The covariances between an asset's illiquidity on one hand, and aggregate illiquidity or returns on the other, are less informative about expected returns. Indeed, these covariances are proportional

29

to cashflow variance Σ_{nn}, which is only a component of $(\Sigma u)_n$. Thus, these covariances proxy for the true pricing factor but imperfectly so.

Corollary 5.2 *In the cross-section of assets, expected returns are proportional to the covariance between returns and aggregate illiquidity. The proportionality coefficient is negative, in both the limit logarithmic $(\gamma = 1, r \to 0)$ and the limit risk-neutral $(\gamma = 0, r \to 0)$ cases.*

The premium associated to the illiquidity risk factor is the expected return per unit of covariance with the factor. We denote it by $\Pi^\Lambda(w_t)$:

$$\frac{\mathrm{E}_t(dR_t)}{dt} \equiv \Pi^\Lambda(w_t)\frac{\mathrm{Cov}_t(d\Lambda_t, dR_t)}{dt}. \tag{5.4}$$

Eqs. (4.2) and (5.3) imply that $\Pi^\Lambda(w_t)$ is related to the common component $C^\Lambda(w_t)$ of assets' covariance with aggregate illiquidity through

$$\Pi^\Lambda(w_t) = \frac{\alpha A(w_t)\left[f(w_t)u^\top\Sigma u + 1\right]}{[\alpha + A(w_t)]\,C^\Lambda(w_t)}. \tag{5.5}$$

The quantities $\Pi^\Lambda(w_t)$ and $C^\Lambda(w_t)$ vary over time in response to changes in arbitrageur wealth. When wealth is low, illiquidity is high and highly sensitive to changes in wealth. Because of this effect, assets' covariance with illiquidity is large and decreases when wealth increases. Conversely, because the premium of the illiquidity risk factor is the expected return per unit of covariance, it is low when wealth is low and increases when wealth increases. For large values of wealth, the premium can decrease again because the decrease in expected returns can dominate the decrease in covariance. Proposition 5.2 derives these results in the limit when r goes to zero, and Figure 6 illustrates them in a numerical example.

Proposition 5.2 *In both the limit logarithmic $(\gamma = 1, r \to 0)$ and the limit risk-neutral $(\gamma = 0, r \to 0)$ cases:*

(i) *The common component $C^\Lambda(w_t) < 0$ of assets' covariance with aggregate illiquidity converges to minus infinity when arbitrageur wealth w_t goes to zero. It remains negative when wealth reaches $\bar{w}$ in the limit risk-neutral case, and converges to zero when wealth goes to infinity in the limit logarithmic case.*

30

(ii) *The premium $\Pi^\Lambda(w_t) < 0$ of the illiquidity risk factor converges to zero when arbitrageur wealth w_t goes to zero. In the limit risk-neutral case, it is inverse-hump shaped in wealth and reaches zero when wealth reaches $\bar{w}$. In the limit logarithmic case, it converges to minus infinity when wealth goes to infinity.*
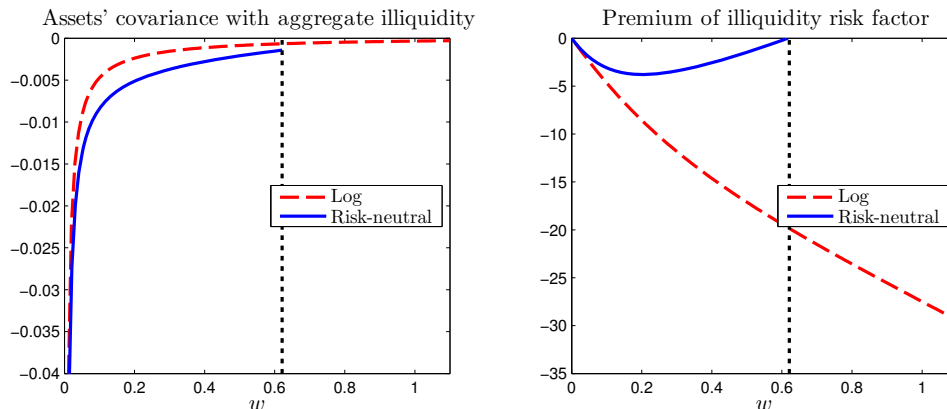

Figure 6: Assets' covariance with aggregate illiquidity (left panel) and premium of illiquidity risk factor (right panel) as a function of arbitrageur wealth, in the logarithmic case (dashed lines) and the risk-neutral case (solid lines). The premium of the illiquidity risk factor is the expected return per unit of covariance. Parameter values are $\alpha = 2$, $\sqrt{u^\top \Sigma u} = 15\%$, $\rho = 4\%$, $r = 2\%$, $N = 2$ symmetric assets with independent cashflows, and $\Sigma_{11} = \Sigma_{22} = 10\%$.

6 Positive supply and infinitely lived CARA hedgers

Until this point we derived our results under the assumption that the supply of risky assets are zero and assumed that hedgers derive utility from the instantaneous changes in wealth. In this section we show that the main insights of our analysis remain intact in variants where long-lived assets are in positive supply and hedgers live forever deriving utility from intertemporal consumption.

Before we analyze the joint effect of hedgers with objective (2.3) and positive supply, let us consider first the effect of positive supply in isolation. Positive s makes the model more directly applicable to stocks and bonds, and to the empirical findings on priced liquidity factors in those markets. Introducing positive supply preserves the basic structure of the equilibrium, with arbitrageur wealth as the only state variable. The following proposition shows the resulting expression for asset prices and expected returns. The rest of the equilibrium objects we spell out in the Appendix.

31

Proposition 6.1 *When the supply vector of risky assets, s, is positive, then asset prices are given by*

$$S(w_t) = \frac{\bar{D} - \alpha\Sigma\left(u + s\right)}{r} + g(w_t)\Sigma\left(u + s\right),\tag{6.1}$$

where $q\left(w_t\right), q_1\left(w_t\right)$ defined as in (3.6) and (3.7) and, along with $g(w_t)$ solve the system of ODEs given in the Appendix. Also, in equilibrium

$$\frac{\mathrm{E}_t(dR_t)}{dt} = \frac{A\left(w_t\right)\alpha\left(f\left(w_t\right)\left(u + s\right)^\top\Sigma\left(u + s\right) + 1\right)\Sigma\left(u + s\right)}{A\left(w_t\right) + \alpha - \alpha g'\left(w\right)\left(u + s\right)^\top\Sigma^\top s}\tag{6.2}$$

Comparing expression (4.11) for expected returns in the baseline case to (6.2) demonstrates the main effects of introducing positive supply. First, u is replaced by $s + u$, because total supply $s + u$ matters for asset prices, and not its breakdown into the supply u coming from hedgers and s coming from asset issuers. Second, because expected returns have to compensate for the risk of holding the total supply, the covariance of the value of the non-tradable part related to u, and the endogenous value of the tradable part related to s will influence expected returns, hence the last term in the denominator.

Suppose that, additionally to positive supply and long-lived assets, hedgers live forever and maximize the objective function (2.3). the following proposition shows the resulting expression for asset prices and expected returns.

Proposition 6.2 *If (3.6)-(3.7) are the value functions of arbitrageurs and*

$$V^h(v_t, w_t) = -e^{-[r\alpha v_t + F(w_t)]},\tag{6.3}$$

is the value function of hedgers where $F(w_t)$ is a function of arbitrageur wealth.and the price is given by

$$S(w_t) = \frac{\bar{D} - r\alpha\Sigma\left(u + s\right)}{r} + g(w_t)\Sigma\left(u + s\right),\tag{6.4}$$

, then the functions $q\left(w_t\right), q_1\left(w_t\right), g\left(w_t\right), F\left(w_t\right)$ must solve the system of ODEs given in the Appendix. Also, expected returns follow

$$\frac{\mathrm{E}_t(dR_t)}{dt} = \frac{A\left(w_t\right)r\alpha}{A\left(w_t\right) - F'\left(w_t\right) + r\alpha - r\alpha g'\left(w\right)\left(u + s\right)^\top\Sigma^\top s}\left(f\left(w_t\right)\left(u + s\right)^\top\Sigma\left(u + s\right) + 1\right)\Sigma\left(u + s\right)$$

$$\tag{6.5}$$

In this case, hedgers' risk aversion reflects not only their myopic demand but also intertemporal hedging demand. This yields an effective risk aversion, $r\alpha - F'(w_t)$, analogous to that of arbitrageurs. This risk aversion depends on arbitrageurs wealth, because the dynamics of prices depends on arbitrageurs wealth. Note that, because of the properties of CARA utility, hedgers' effective risk aversion does not depend on hedgers' wealth v_t. The function $F(w_t)$ can be characterized by an ODE which must be solved numerically simultaneously with arbitrageurs value function and the asset prices. While expression (6.4) determining $S(w_t)$ looks very similar to (6.1), the fact that effective risk aversion of hedgers is changed to $r\alpha - F'(w_t)$ is visible from the additional r in the numerator of the first term and it affects $g(w_t)$ through $F(w_t)$ (as it is shown in the Appendix).

The next proposition states that none of these extensions change the result that expected returns are proportional to the covariance of asset prices and aggregate illiquidity.

Proposition 6.3 *In the cross-section of assets, expected returns are proportional to the covariance between returns and aggregate illiquidity even when $s > 0$ and even with long-term CARA hedgers.*

The proposition holds because of the same argument as we have seen in the baseline model.

On Figure 7 we compare the main equilibrium objects with the baseline model (dotted), the case with positive supply (dashed) and the case with both positive supply and long-lived hedgers (solid). The parameters of the baseline case are as before. In the other two scenarios $s = 0.5$ and $\left((u+s)^\top \Sigma (u+s)\right)^{\frac{1}{2}} = 15\%$, $r = 2\%, \rho_h = \rho = 4\%$. While the short-lived hedgers case we keep $\alpha = 2$, we choose $r\alpha = 2$ in the long-lived hedgers economy so all three economies coincide without arbitrageurs. For brevity, we show here the $\gamma = 1$ case only. It is apparent that neither modifications change the major qualitative properties of the equilibrium compared to the baseline model.

7   Concluding Remarks

We develop a dynamic model of liquidity provision, in which hedgers can trade multiple risky assets with arbitrageurs. We compute the equilibrium in closed form when arbitrageurs' utility over consumption is logarithmic or risk-neutral with a non-negativity constraint. Our model provides an explanation for why liquidity varies over time and is a priced risk factor: liquidity decreases following losses by arbitrageurs, assets with volatile cashflows or in high supply by hedgers suffer
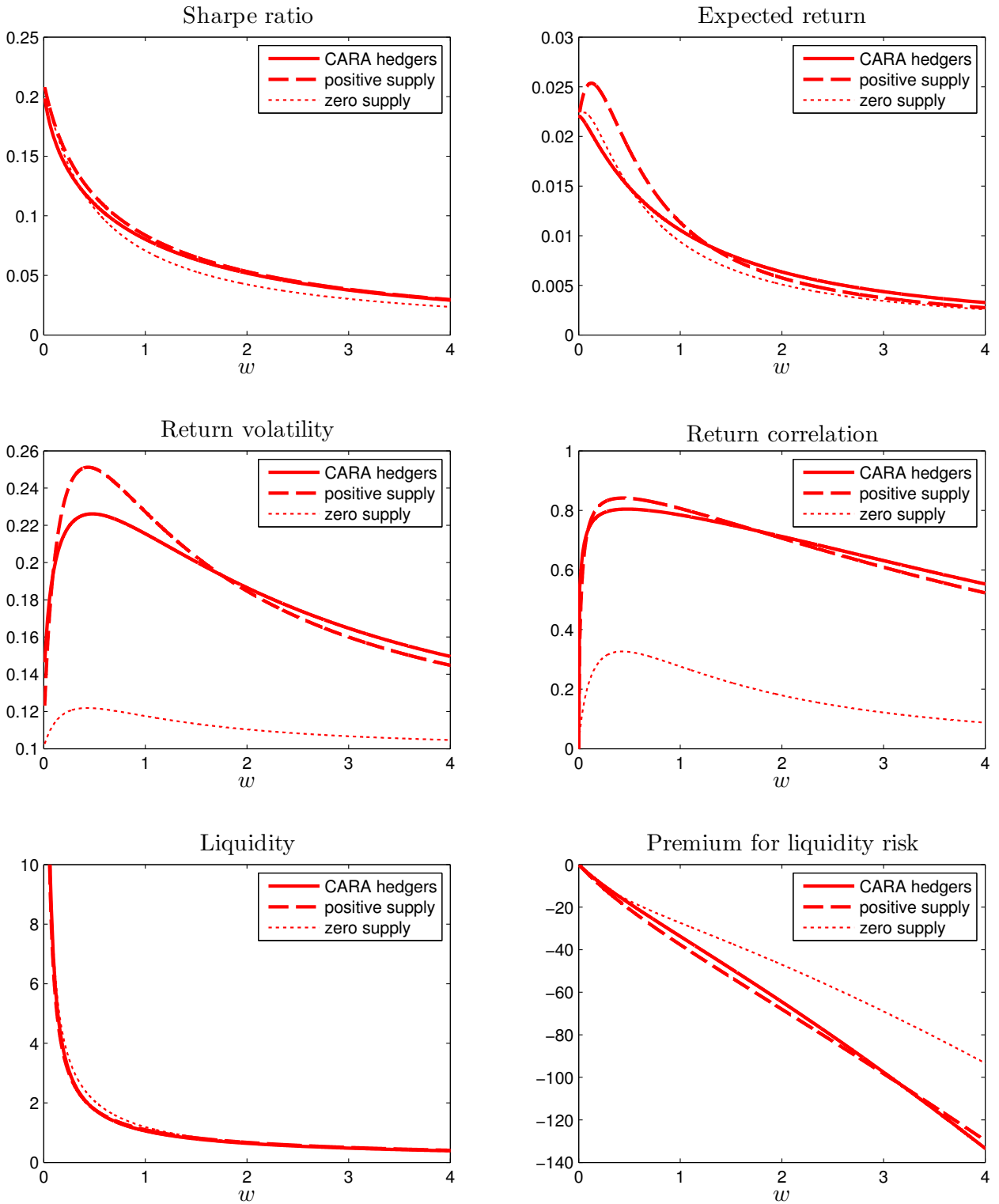
Figure 7: Equilibrium objects with log-arbitrageurs ($\gamma = 1$) under three scenarios. In each scenario there are two risky assets with independent and identically distributed payoffs. The first scenario (dotted) is the baseline case with zero supply. In the second scenario (dashed) hedgers are short-term mean-variance investors, while in the third one (solid) hedgers are infinitely lived CARA investors. Parameter values are $r = 0.02$, $\rho = 0.04$, $\sigma_1 = \sigma_2 = 0.1$, $(u + s)^\top \Sigma (u + s) = 0.0225$, $s = 0.5$ (second and third scenario) and $\alpha = 2$ (first and second scenario) and $r\alpha = 2$ (third scenario).

34

the most from low liquidity, and these assets offer the highest expected returns. Our model also provides a broader framework for analyzing the dynamics of arbitrage capital and its link with asset prices and risk-sharing. Among other results, we show that asset volatilities, correlations, and expected returns are hump-shaped functions of arbitrageur wealth. We also show that when hedgers become more risk averse or asset cashflows become more volatile, arbitrageurs can choose to provide less liquidity even though liquidity provision becomes more profitable. Finally, we characterize the stationary distribution of arbitrageur wealth, and show that it becomes bimodal when hedging needs are strong.

An interesting extension to our set up is to allow the supply u coming from hedgers to be stochastic. A stochastic u gives our measure of illiquidity (5.1) stronger empirical content since that measure is based on changes in u. It is possible to introduce stochastic u and yet preserve much of the tractability of our model provided that we restrict the variance $u^\top \Sigma u$ of hedgers' endowment to remain constant over time. This is because the dynamics of arbitrageur wealth derived in Section 3.2 depend on the N-dimensional vector u only through the one-dimensional statistic $u^\top \Sigma u$. Restricting $u^\top \Sigma u$ to be constant allows for hedger supply to, e.g., increase for some assets provided that it decreases for others. In Kondor and Vayanos (2014) we present a parsimonious specification with stochastic u and constant $u^\top \Sigma u$, and confirm that the dynamics of arbitrageur wealth are identical to those derived in Section 3.2. Eqs. (4.9) and (4.11) for asset prices and expected returns remain valid, although the function $g(w_t)$ cannot be derived in closed form. The results on liquidity risk remain the same. The extension of stochastic u can be combined with that of positive s; tractability is preserved with the additional restriction that $u^\top \Sigma s$ remains constant over time. As future research, we plan to use this extension to understand better whether popular empirical measures of liquidity innovations capture shocks to hedgers' demand or to arbitrageurs wealth. We also plan to assess the quantitative implications of our model.

APPENDIX—Proofs

Proof of Proposition 3.1: Substituting dD_t from (2.1), we can write (3.3) as

$$dv_t = rv_t dt + x_t^\top(\bar{D} - \pi_t)dt + u^\top \bar{D}dt + (x_t + u)^\top \sigma^\top dB_t. \tag{A.1}$$

Substituting dv_t from (A.1) into (2.2), we can write the hedger's maximization problem as

$$\max_{x_t} \left\{ x_t^\top(\bar{D} - \pi_t) - \frac{\alpha}{2}(x_t + u)^\top \Sigma (x_t + u) \right\}. \tag{A.2}$$

The first-order condition is

$$\bar{D} - \pi_t - \alpha\Sigma(x_t + u). \tag{A.3}$$

Solving for x_t, we find (3.4). ∎

Proof of Proposition 3.2: The Bellman equation is

$$\rho V = \max_{\hat{c}_t, \hat{y}_t} \left\{ u(\hat{c}_t) + V_{\hat{w}}\mu_{\hat{w}t} + \frac{1}{2}V_{\hat{w}\hat{w}}\sigma_{\hat{w}t}^\top\sigma_{\hat{w}t} + V_w\mu_{wt} + \frac{1}{2}V_{ww}\sigma_{wt}^\top\sigma_{wt} + V_{\hat{w}w}\sigma_{\hat{w}t}^\top\sigma_{wt} \right\}, \tag{A.4}$$

where $u(\hat{c}_t) = \frac{\hat{c}_t^{1-\gamma}}{1-\gamma}$ for $\gamma \neq 1$ and $u(\hat{c}_t) = \log(\hat{c}_t)$ for $\gamma = 1$, $(\mu_{\hat{w}t}, \sigma_{\hat{w}t})$ are the drift and diffusion of the arbitrageur's own wealth $\hat{w}_t$, and (μ_{wt}, σ_{wt}) are the drift and diffusion of the arbitrageurs' total wealth. Substituting dD_t from (2.1), we can write (3.5) as

$$dw_t = (rw_t - c_t)dt + y_t^\top(\bar{D} - \pi_t)dt + y_t^\top \sigma^\top dB_t. \tag{A.5}$$

Eq. (A.5) written for own wealth implies that

$$\mu_{\hat{w}t} = r\hat{w}_t - \hat{c}_t + \hat{y}_t^\top(\bar{D} - \pi_t), \tag{A.6}$$

$$\sigma_{\hat{w}t} = \sigma\hat{y}_t, \tag{A.7}$$

and written for total wealth implies that

$$\mu_{wt} = rw_t - c_t + y_t^\top(\bar{D} - \pi_t), \tag{A.8}$$

$$\sigma_{wt} = \sigma y_t. \tag{A.9}$$

36

When $\gamma \neq 1$, we substitute (3.6) and (A.6)-(A.9) into (A.4) to write it as

$$\rho q(w_t)\frac{\hat{w}_t^{1-\gamma}}{1-\gamma} = \max_{\hat{c}_t,\hat{y}_t}\left\{\frac{\hat{c}_t^{1-\gamma}}{1-\gamma} + q(w_t)\hat{w}_t^{-\gamma}\left(r\hat{w}_t - \hat{c}_t + \hat{y}_t^\top(\bar{D} - \pi_t)\right) - \frac{1}{2}q(w_t)\gamma\hat{w}_t^{-\gamma-1}\hat{y}_t^\top\Sigma\hat{y}_t\right.$$

$$\left. + q'(w_t)\frac{\hat{w}_t^{1-\gamma}}{1-\gamma}\left(rw_t - c_t + y_t^\top(\bar{D} - \pi_t)\right) + \frac{1}{2}q''(w_t)\frac{\hat{w}_t^{1-\gamma}}{1-\gamma}y_t^\top\Sigma y_t + q'(w_t)\hat{w}_t^{-\gamma}\hat{y}_t^\top\Sigma y_t\right\}. \quad \text{(A.10)}$$

The first-order conditions with respect to $\hat{c}_t$ and $\hat{y}_t$ are (3.8) and (3.9), respectively.

When $\gamma = 1$, we substitute (3.7), (A.6), (A.7), (A.8), and (A.9) into (A.4) to write it as

$$\rho\left(\frac{1}{\rho}\log(\hat{w}_t) + q_1(w_t)\right) = \max_{\hat{c}_t,\hat{y}_t}\left\{\log(\hat{c}_t) + \frac{1}{\rho\hat{w}_t}\left(r\hat{w}_t - \hat{c}_t + \hat{y}_t^\top(\bar{D} - \pi_t)\right) - \frac{1}{2\rho\hat{w}_t^2}\hat{y}_t^\top\Sigma\hat{y}_t\right.$$

$$\left. + q_1'(w_t)\left(rw_t - c_t + y_t^\top(\bar{D} - \pi_t)\right) + \frac{1}{2}q_1''(w_t)y_t^\top\Sigma y_t\right\}. \quad \text{(A.11)}$$

The first-order conditions with respect to $\hat{c}_t$ and $\hat{y}_t$ are (3.8) and (3.9) for $q(w_t) = \frac{1}{\rho}$. ∎

Proof of Proposition 3.3: Since in equilibrium $\hat{c}_t = c_t$ and $\hat{w}_t = w_t$, (3.8) implies that

$$c_t = q(w_t)^{-\frac{1}{\gamma}}w_t. \quad \text{(A.12)}$$

Substituting (3.13) and (3.14) into (3.9) and using the definition of $A(w_t)$ from (3.11), we find

$$\hat{y}_t = \frac{\alpha\hat{w}_t}{(\alpha + A(w_t))w_t}u. \quad \text{(A.13)}$$

When $\gamma \neq 1$, we substitute (3.8), (3.13), (3.14), (A.12), and (A.13) into (A.10). The terms in $\hat{w}_t$ cancel, and the resulting equation is

$$\rho q(w_t) = q(w_t)^{1-\frac{1}{\gamma}} + \left(q'(w_t) + \frac{q(w_t)(1-\gamma)}{w_t}\right)\left(rw_t - q(w_t)^{-\frac{1}{\gamma}}w_t + \frac{\alpha^2 A(w_t)}{(\alpha + A(w_t))^2}u^\top\Sigma u\right)$$

$$+ \frac{1}{2}\left(q''(w_t) - \frac{q(w_t)\gamma(1-\gamma)}{w_t^2} + \frac{2q'(w_t)(1-\gamma)}{w_t}\right)\frac{\alpha^2}{(\alpha + A(w_t))^2}u^\top\Sigma u. \quad \text{(A.14)}$$

Using the definition of $A(w_t)$ and rearranging, we find (3.16).

When $\gamma = 1$, we substitute (3.8), (3.13), (3.14), (A.12), and (A.13) into (A.11), setting $q(w_t) = \frac{1}{\rho}$. (Note that this value of $q(w_t)$ solves (3.16) for $\gamma = 1$.) The terms in $\hat{w}_t$ cancel, and the resulting

37

equation is

$$\rho q_1(w_t) = \log(\rho) + \left(q_1'(w_t) + \frac{1}{\rho w_t}\right)\left(rw_t - \rho w_t + \frac{\alpha^2 A(w_t)}{(\alpha + A(w_t))^2}u^\top \Sigma u\right)$$

$$+ \frac{1}{2}\left(q_1''(w_t) - \frac{1}{\rho w_t^2}\right)\frac{\alpha^2}{(\alpha + A(w_t))^2}u^\top \Sigma u. \tag{A.15}$$

Using the definition of $A(w_t)$ and rearranging, we find (3.17). ∎

Lemma A.1 recalls some useful properties of the cotangent function.

Lemma A.1 *The function $x\cot(x)$ is decreasing for $x \in [0, \frac{\pi}{2}]$. Its asymptotic behavior for x close to zero is*

$$x\cot(x) = 1 - \frac{x^2}{3} + o(x^2). \tag{A.16}$$

Proof: Differentiating $x\cot(x)$ with respect to x, we find

$$\frac{d}{dx}[x\cot(x)] = \cot(x) - x\left[1 + \cot^2(x)\right]$$

$$= \cot(x)\left[1 - \frac{x}{\sin(x)\cos(x)}\right]$$

$$= \cot(x)\left[1 - \frac{2x}{\sin(2x)}\right]. \tag{A.17}$$

The function $x - \sin(x)$ is equal to zero for $x = 0$, and its derivative $1 - \cos(x)$ is positive for $x \in (0, \pi)$. Therefore, $x > \sin(x)$ for $x \in (0, \pi)$. Since, in addition, $\sin(x) > 0$ for $x \in (0, \pi)$ and $\cot(x) > 0$ for $x \in (0, \frac{\pi}{2})$, (A.17) is negative for $x \in (0, \frac{\pi}{2})$ and so $x\cot(x)$ is decreasing. Using the asymptotic behavior of $\sin(x)$ and $\cos(x)$ for x close to zero, we find

$$\cot(x) = \frac{\cos(x)}{\sin(x)} = \frac{1 - \frac{x^2}{2} + o(x^2)}{x - \frac{x^3}{6} + o(x^3)} = \frac{1}{x}\left(1 - \frac{x^2}{3} + o(x^2)\right),$$

which implies (A.16). ∎

Proof of Proposition 3.4: For $\gamma = 0$ and $w_t < \bar{w}$, (3.16) becomes

$$(\rho - r)q = rq'w + \frac{1}{2}\left(q'' - \frac{2q'^2}{q}\right)\frac{\alpha^2}{\left(\alpha - \frac{q'}{q}\right)^2}u^\top \Sigma u. \tag{A.18}$$

38

Dividing both sides by $q(w_t)$, and noting that $A(w_t) = -\frac{q'(w_t)}{q(w_t)}$ for $\gamma = 0$, we can write (A.18) as

$$\rho - r = -rAw - \frac{1}{2}\left(A' + A^2\right)\frac{\alpha^2}{(\alpha + A)^2}u^\top \Sigma u. \tag{A.19}$$

Eq. (A.19) is a first-order ODE in the function $A(w_t)$. It must be solved with the boundary condition $\lim_{w_t \to 0} A(w_t) = \infty$. This is because when w_t goes to zero, arbitrageurs' position y_t in the risky assets should go to zero so that w_t remains non-negative, and y_t is given by (3.14).

In the limit when r goes to zero, (A.19) becomes

$$\rho = -\frac{1}{2}\left(A' + A^2\right)\frac{\alpha^2}{(\alpha + A)^2}u^\top \Sigma u$$

$$\Leftrightarrow 1 = -\left(A' + A^2\right)\frac{z}{(\alpha + A)^2} \tag{A.20}$$

$$\Leftrightarrow -\frac{zA'}{zA^2 + (\alpha + A)^2} = 1$$

$$\Leftrightarrow -\frac{\frac{zA'}{1+z}}{\left(A + \frac{\alpha}{1+z}\right)^2 + \frac{z\alpha^2}{(1+z)^2}} = 1, \tag{A.21}$$

where (A.20) follows from the definition of z. Setting

$$\hat{A}(w_t) \equiv \frac{1+z}{\alpha\sqrt{z}}\left(A(w_t) + \frac{\alpha}{1+z}\right),$$

we can write (A.21) as

$$-\frac{\hat{A}'}{\hat{A}^2 + 1} = \frac{\alpha}{\sqrt{z}}. \tag{A.22}$$

Eq. (A.22) integrates to

$$\operatorname{arccot}\left(\hat{A}(w_t)\right) - \operatorname{arccot}\left(\hat{A}(0)\right) = \frac{\alpha w_t}{\sqrt{z}}. \tag{A.23}$$

The boundary condition $\lim_{w_t \to 0} A(w_t) = \infty$ implies $\lim_{w_t \to 0} \hat{A}(w_t) = \infty$ and hence

$$\operatorname{arccot}\left(\hat{A}(0)\right) = 0.$$

Substituting into (A.23), we find

$$\hat{A}(w_t) = \cot\left(\frac{\alpha w_t}{\sqrt{z}}\right). \tag{A.24}$$

Eq. (A.24) and the definition of $\hat{A}(w_t)$ imply that $A(w_t)$ is given by (3.20) for $w_t < \bar{w}$. Since $q(w_t) = 1$ for $w_t \geq \bar{w}$, $A(w_t) = -\frac{q'(w_t)}{q(w_t)}$ implies that $A(w_t) = 0$ for $w_t \geq \bar{w}$. Smooth-pasting implies that $A(w_t)$ given by (3.20) must be equal to zero for $w_t = \bar{w}$. This yields (3.21).

To determine $q(w_t)$, we solve

$$\frac{q'}{q} = -A \tag{A.25}$$

with the boundary condition $q(\bar{w}) = 1$. Eq. (A.25) integrates to

$$\log q(w_t) - \log q(\bar{w}) = \int_{w_t}^{\bar{w}} A(w)dw$$

$$\Rightarrow \log q(w_t) = \int_{w_t}^{\bar{w}} A(w)dw, \tag{A.26}$$

where the second step follows from the boundary condition. Substituting $A(w_t)$ from (3.20) into (A.26) and integrating, we find (3.22).

We finally show that $A(w_t)$ is decreasing and convex, converges to ∞ when w_t goes to zero, and is smaller than $\frac{1}{w_t}$. Since the right-hand side of (3.21) is positive, $\frac{\alpha\bar{w}}{\sqrt{z}} < \frac{\pi}{2}$. Since $\cot(x)$ is decreasing for $x \in (0, \frac{\pi}{2})$, (3.20) implies that $A(w_t)$ is decreasing for $w_t \in (0, \bar{w}]$. Differentiating (3.20) with respect to w_t yields

$$A'(w_t) = -\frac{\alpha^2}{1+z}\left(1 + \cot^2\left(\frac{\alpha w_t}{\sqrt{z}}\right)\right). \tag{A.27}$$

Since $\cot(x)$ is positive and decreasing for $x \in (0, \frac{\pi}{2})$, $A'(w_t)$ is increasing for $w_t \in (0, \bar{w}]$. Therefore, $A(w_t)$ is convex. Since $\cot(x)$ converges to ∞ when x goes to zero, (3.20) implies that $A(w_t)$ converges to ∞ when w_t goes to zero. Since the function $x \cot(x)$ is decreasing, it is smaller than one, its limit when x goes to zero (Lemma A.1). Therefore, (3.20) implies that

$$w_t A(w_t) = \frac{\alpha w_t}{1+z}\left(\sqrt{z}\cot\left(\frac{\alpha w_t}{\sqrt{z}}\right) - 1\right) < \frac{\alpha\sqrt{z}w_t}{1+z}\cot\left(\frac{\alpha w_t}{\sqrt{z}}\right) < \frac{z}{1+z} < 1,$$

40

and hence $A(w_t) < \frac{1}{w_t}$. ∎

Proof of Proposition 3.5: In both the logarithmic and the limit risk-neutral cases, $A(w_t)$ is decreasing in w_t. Part (i) follows from this property, (3.1) and (3.13). Part (ii) follows from the same property and (3.14). Part (iii) follows from the same property and because (3.13) implies that the market prices of risk are given by

$$\left(\sigma^\top\right)^{-1}(\bar{D} - \pi_t) = \frac{\alpha A(w_t)}{\alpha + A(w_t)}\sigma u.$$

Part (iv) follows from the same property and (3.15). ∎

Proof of Proposition 3.6: Substituting (3.13), (3.14), and (A.12) into (A.5) we can write the dynamics of arbitrageur wealth w_t as

$$dw_t = \mu_{wt}dt + \sigma_{wt}^\top dB_t, \tag{A.28}$$

where

$$\mu_{wt} \equiv \left(r - q(w_t)^{-\frac{1}{\gamma}}\right)w_t + \frac{\alpha^2 A(w_t)}{(\alpha + A(w_t))^2}u^\top \Sigma u, \tag{A.29}$$

$$\sigma_{wt} \equiv \frac{\alpha}{\alpha + A(w_t)}\sigma u. \tag{A.30}$$

We first determine the stationary distribution in the limit risk-neutral case. Wealth evolves in $(0, \bar{w})$, with an upper reflecting barrier at $\bar{w}$. Since the consumption rate $q(w_t)^{-\frac{1}{\gamma}}$ is equal to zero in $(0, \bar{w})$ and r is equal to zero in the limit risk-neutral case, we can write the drift (A.29) as

$$\mu_{wt} = \frac{\alpha^2 A(w_t)}{(\alpha + A(w_t))^2}u^\top \Sigma u. \tag{A.31}$$

If the stationary distribution has density $d(w_t)$, then $d(w_t)$ satisfies the ODE

$$-(\mu_w d)' + \frac{1}{2}(\sigma_w^T \sigma_w d)'' = 0 \tag{A.32}$$

over $(0, \bar{w})$, and the boundary condition

$$-\mu_w d + \frac{1}{2}(\sigma_w^T \sigma_w d)' = 0 \tag{A.33}$$

41

at the reflecting barrier $\bar{w}$. Integrating (A.32) using (A.33) yields the ODE

$$-\mu_w d + \frac{1}{2}(\sigma_w^T \sigma_w d)' = 0. \tag{A.34}$$

Setting $D(w_t) \equiv \sigma_{wt}^T \sigma_{wt} d(w_t)$, we can write (A.34) as

$$\frac{D'}{D} = \frac{2\mu_w}{\sigma_w^T \sigma_w}. \tag{A.35}$$

Eq. (A.35) integrates to

$$D(w_t) = D(\bar{w}) \exp\left(-\int_{w_t}^{\bar{w}} \frac{2\mu_w}{\sigma_w^T \sigma_w} dw\right),$$

yielding

$$d(w_t) = D(\bar{w}) \frac{\exp\left(-\int_{w_t}^{\bar{w}} \frac{2\mu_w}{\sigma_w^T \sigma_w} dw\right)}{\sigma_{wt}^T \sigma_{wt}}. \tag{A.36}$$

We can determine the multiplicative constant $D(\bar{w})$ by the requirement that $d(w_t)$ must integrate to one, i.e.,

$$\int_0^{\bar{w}} d(w_t) dw_t = D(\bar{w}) \int_0^{\bar{w}} \frac{\exp\left(-\int_{w_t}^{\bar{w}} \frac{2\mu_w}{\sigma_w^T \sigma_w} dw\right)}{\sigma_{wt}^T \sigma_{wt}} dw_t = 1. \tag{A.37}$$

Eq. (A.37) determines a positive $D(\bar{w})$, and hence a positive $d(w_t)$, if the integral multiplying $D(\bar{w})$ is finite. If the integral is infinite, then (A.37) implies that $D(\bar{w}) = 0$, and the stationary distribution does not have a density but is concentrated at zero. The integral multiplying $D(\bar{w})$ is infinite when the integrand converges to infinity at a fast enough rate when w_t goes to zero.

Substituting μ_{wt} and σ_{wt} from (A.31) and (A.30), respectively, into (A.36), we find

$$d(w_t) = \frac{D(\bar{w})}{\alpha^2 u^\top \Sigma u}(\alpha + A(w_t))^2 \exp\left(-2\int_{w_t}^{\bar{w}} A(w) dw\right)$$

$$= \frac{D(\bar{w})}{\alpha^2 u^\top \Sigma u}\left(\frac{\alpha + A(w_t)}{q(w_t)}\right)^2, \tag{A.38}$$

where the second step follows from (A.26). Eqs. (A.37) and (A.38) imply (3.24). Eqs. (3.20) and (3.22), and Lemma A.1, imply that when w_t is close to zero,

$$\left(\frac{\alpha + A(w_t)}{q(w_t)}\right)^2 \approx \Gamma\left(\frac{\frac{1}{w_t}}{\exp\left(-\frac{z}{1+z}\log(w_t)\right)}\right)^2 = \Gamma\left(\frac{w_t^{\frac{z}{1+z}}}{w_t}\right)^2 = \Gamma w_t^{-\frac{2}{1+z}},$$

where Γ is a positive constant. Therefore, the integral multiplying $D(\bar{w})$ in (A.37) is finite when

$$\frac{2}{1+z} < 1 \Leftrightarrow z > 1.$$

We next determine the stationary distribution in the logarithmic case. Wealth evolves in $(0, \infty)$. Noting that $c_t = \rho w_t$ and $A(w_t) = \frac{1}{w_t}$, we can write the drift (A.29) and the diffusion (A.30) as

$$\mu_{wt} \equiv (r - \rho)w_t + \frac{\alpha^2 w_t}{(\alpha w_t + 1)^2} u^\top \Sigma u, \tag{A.39}$$

$$\sigma_{wt} \equiv \frac{\alpha w_t}{\alpha w_t + 1} \sigma u, \tag{A.40}$$

respectively. In the logarithmic case there is no reflecting barrier, but (A.34) still holds. Intuitively, (A.34) holds for any reflecting barrier, and the effect of a reflecting barrier on the stationary distribution converges to zero when the barrier goes to infinity. To compute the density $d(w_t)$ of the stationary distribution, we thus need to integrate (A.35). Integrating between an arbitrary value $\bar{w}_0$ and w_t, we find

$$d(w_t) = D(\bar{w}_0) \frac{\exp\left(-\int_{w_t}^{\bar{w}_0} \frac{2\mu_w}{\sigma_w^T \sigma_w} dw\right)}{\sigma_{wt}^T \sigma_{wt}}. \tag{A.41}$$

We can determine the multiplicative constant $D(\bar{w}_0)$ by the requirement that $d(w_t)$ must integrate to one, i.e.,

$$\int_0^\infty d(w_t) dw_t = D(\bar{w}_0) \int_0^\infty \frac{\exp\left(-\int_{w_t}^{\bar{w}_0} \frac{2\mu_w}{\sigma_w^T \sigma_w} dw\right)}{\sigma_{wt}^T \sigma_{wt}} dw_t = 1. \tag{A.42}$$

Substituting μ_{wt} and σ_{wt} from (A.39) and (A.40) into (A.41), we find

$$d(w_t) = D(\bar{w}_0) \frac{\exp\left(-\int_{w_t}^{\bar{w}_0} \left(-\frac{w\left(\alpha+\frac{1}{w}\right)^2}{z} + \frac{2}{w}\right) dw\right)}{\frac{\alpha^2 u^\top \Sigma u}{\left(\alpha+\frac{1}{w_t}\right)^2}}$$

$$= C(\bar{w}_0) \frac{\exp\left(-\frac{1}{z}\left(\frac{1}{2}\alpha^2 w_t^2 + 2\alpha w_t + \log(w_t)\right) + 2\log(w_t)\right)}{\frac{\alpha^2 u^\top \Sigma u w_t^2}{(\alpha w_t + 1)^2}}$$

$$= \frac{C(\bar{w}_0)}{\alpha^2 u^\top \Sigma u} (\alpha w_t + 1)^2 w_t^{-\frac{1}{z}} \exp\left(-\frac{1}{2z}\left(\alpha^2 w_t^2 + 4\alpha w_t\right)\right), \tag{A.43}$$

43

where

$$C(\bar{w}_0) \equiv D(\bar{w}_0) \exp\left(\frac{1}{z}\left(\frac{1}{2}\alpha^2\bar{w}_0^2 + 2\alpha\bar{w}_0 + \log(\bar{w}_0)\right) - 2\log(\bar{w}_0)\right).$$

Eqs. (A.42) and (A.43) imply (3.23). If $z < 0$, then the integral multiplying $D(\bar{w}_0)$ is infinite because of the behavior of the integrand when z goes to ∞. If $z > 0$, then the integral can be infinite because of the behavior of the integrand when w_t is close to zero. Since

$$(\alpha w_t + 1)^2 w_t^{-\frac{1}{z}} \exp\left(-\frac{1}{2z}\left(\alpha^2 w_t^2 + 4\alpha w_t\right)\right) \approx w_t^{-\frac{1}{z}}$$

when w_t is close to zero, the integral multiplying $D(\bar{w})$ in (A.42) is finite when $z > 1$. ∎

Proof of Proposition 3.7: Eq. (3.23) implies that in the logarithmic case, the derivative of $d(w_t)$ with respect to w_t has the same sign as the derivative of

$$(\alpha w_t + 1)^2 w_t^{-\frac{1}{z}} \exp\left(-\frac{1}{2z}\left(\alpha^2 w_t^2 + 4\alpha w_t\right)\right).$$

The latter derivative is

$$\frac{1}{z}(\alpha w_t + 1)w_t^{-\frac{1}{z}-1} \exp\left(-\frac{1}{2z}\left(\alpha^2 w_t^2 + 4\alpha w_t\right)\right)[2\alpha z w_t - (\alpha w_t + 1) - (\alpha w_t + 1)\alpha w_t(\alpha w_t + 2)]$$

and has the same sign as

$$-\left[(\alpha w_t)^3 + 3(\alpha w_t)^2 + (3 - 2z)\alpha w_t + 1\right].$$

The function

$$F(x) \equiv x^3 + 3x^2 + (3 - 2z)x + 1$$

is equal to 1 for $x = 0$, and its derivative with respect to x is

$$F'(x) = 3x^2 + 6x + (3 - 2z).$$

If $z < \frac{3}{2}$, then $F'(x) > 0$ for all $x > 0$, and hence $F(x) > 0$ for all $x > 0$. If $z > \frac{3}{2}$, then $F'(x)$ has the positive root

$$x_1' \equiv -1 + \sqrt{\frac{2z}{3}},$$

and is negative for $0 < x < x_1'$ and positive for $x > x_1'$. Therefore, if $F(x_1') > 0$ then $F(x) > 0$ for all $x > 0$, and if $F(x_1') < 0$ then $F(x)$ has two positive roots $x_1 < x_1' < x_2$ and is positive outside the roots and negative inside. Since

$$F(x_1') = \left(-1 + \sqrt{\frac{2z}{3}}\right)^3 + 3\left(-1 + \sqrt{\frac{2z}{3}}\right)^2 + (3 - 2z)\left(-1 + \sqrt{\frac{2z}{3}}\right) + 1 = \frac{2z}{3}\left(3 - 2\sqrt{\frac{2z}{3}}\right),$$

$F(x_1')$ is positive if

$$3 - 2\sqrt{\frac{2z}{3}} > 0 \equiv z < \frac{27}{8},$$

and is negative if $z > \frac{27}{8}$. Therefore, if $z < \frac{27}{8}$ then the derivative of $d(w_t)$ is negative, and if $z > \frac{27}{8}$ then the derivative of $d(w_t)$ is negative for $w_t \in (0, \bar{w}_1) \cup (\bar{w}_2, \infty)$ and positive for $w_t \in (\bar{w}_1, \bar{w}_2)$, where $\bar{w}_i \equiv \frac{x_i}{\alpha}$ for $i = 1, 2$. This proves Part (i).

Eq. (3.24) implies that in the limit risk-neutral case, the derivative of $d(w_t)$ with respect to w_t has the same sign as the derivative of $\frac{\alpha + A(w_t)}{q(w_t)}$. The latter derivative is

$$\frac{d}{dw_t}\left(\frac{\alpha + A(w_t)}{q(w_t)}\right) = \frac{A'(w_t)q(w_t) - (\alpha + A(w_t))q'(w_t)}{q(w_t)^2}$$

$$= \frac{A'(w_t) + A(w_t)(\alpha + A(w_t))}{q(w_t)}$$

$$= \frac{-\frac{(\alpha + A(w_t))^2}{z} + \alpha A(w_t)}{q(w_t)},$$

where the second step follows from (A.25) and the third from (A.20). Therefore, the derivative of $d(w_t)$ with respect to w_t has the same sign as

$$-\left[\alpha^2 + A(w_t)^2 - (z - 2)\alpha A(w_t)\right]. \tag{A.44}$$

The term in square brackets is a quadratic function of $A(w_t)$ and is always positive if

$$(z - 2)^2 - 4 < 0 \Leftrightarrow z < 4.$$

Therefore, if $z < 4$, then the derivative of $d(w_t)$ is negative. If $z > 4$, then the quadratic function has two positive roots, given by

$$\frac{z - 2 \pm \sqrt{z(z - 4)}}{2},$$

45

and is positive outside the roots and negative inside. We define the thresholds $\bar{w}_1$ and $\bar{w}_2$ such that $A(\bar{w}_1)$ is equal to the smaller root and $A(\bar{w}_2)$ is equal to the larger root. Since $A(w_t)$ decreases from infinity to zero when w increases from zero to $\bar{w}$, the thresholds $\bar{w}_1$ and $\bar{w}_2$ are uniquely defined and satisfy $0 < \bar{w}_1 < \bar{w}_2 < \bar{w}$. Moreover, the derivative of $d(w_t)$ is negative for $w_t \in (0, \bar{w}_1) \cup (\bar{w}_2, \bar{w})$ and positive for $w_t \in (\bar{w}_1, \bar{w}_2)$. This proves Part (ii).

The density $d(w_t)$ shifts to the right in the monotone likelihood ratio sense when a parameter θ increases if

$$\frac{\partial^2 \log\left(d(w_t, \theta)\right)}{\partial \theta \partial w_t} > 0. \tag{A.45}$$

Using (3.23), we find that in the logarithmic case,

$$\frac{\partial \log(d(w_t))}{\partial w_t} = \frac{2\alpha}{\alpha w_t + 1} - \frac{1}{z w_t} - \frac{1}{z}\left(\alpha^2 w_t + 2\alpha\right). \tag{A.46}$$

An increase in α (which also affects z) raises the right-hand side of (A.46). Therefore, $d(w_t)$ satisfies (A.45) with respect to α. An increase in z also raises the right-hand side of (A.46). Therefore, $d(w_t)$ satisfies (A.45) with respect to $u^\top \Sigma u$. Using (3.24), we find that in the limit risk-neutral case,

$$\begin{aligned}
\frac{\partial \log(d(w_t))}{\partial w_t} &= \frac{2A'(w_t)}{\alpha + A(w_t)} - \frac{2q'(w_t)}{q(w_t)} \\
&= \frac{2A'(w_t)}{\alpha + A(w_t)} + 2A(w_t),
\end{aligned} \tag{A.47}$$

where the second step follows from (A.25). Eqs. (3.20) and (A.27) imply that

$$\frac{A'(w_t)}{\alpha + A(w_t)} = -\frac{1 + \cot^2\left(\frac{\alpha w_t}{\sqrt{z}}\right)}{\sqrt{z}\left(\cot\left(\frac{\alpha w_t}{\sqrt{z}}\right) + \sqrt{z}\right)}. \tag{A.48}$$

An increase in α (which also affects z) raises the right-hand side of (A.48). Since it also raises $A(w_t)$ (Part (i) of Lemma 3.4), (A.47) implies that $d(w_t)$ satisfies (A.45) with respect to α. Differentiating (A.48) with respect to $\sqrt{z}$, we find

$$\frac{\partial}{\partial \sqrt{z}}\left(\frac{A'(w_t)}{\alpha + A(w_t)}\right) = \frac{\left(1 + \cot^2\left(\frac{\alpha w_t}{\sqrt{z}}\right)\right)\left[\left(2\sqrt{z} + \cot\left(\frac{\alpha w_t}{\sqrt{z}}\right)\right)\left(1 - \frac{\alpha w_t}{\sqrt{z}}\cot\left(\frac{\alpha w_t}{\sqrt{z}}\right)\right) + \frac{\alpha w_t}{\sqrt{z}}\right]}{z\left(\cot\left(\frac{\alpha w_t}{\sqrt{z}}\right) + \sqrt{z}\right)^2}.$$

46

Since the function $x \cot(x)$ is smaller than one (Lemma A.1), the term in square brackets is positive. Therefore, an increase in z raises the right-hand side of (A.48). Since it also raises $A(w_t)$ (Part (iii) of Lemma 3.4), (A.47) implies that $d(w_t)$ satisfies (A.45) with respect to $u^\top \Sigma u$. This proves Part (iii). ∎

Lemma A.2 shows some useful properties of $A(w_t)$.

Lemma A.2 *Suppose that* $\gamma = r = 0$.

(i) *An increase in* α *raises* $A(w_t)$.

(ii) *An increase in* α *raises* $\frac{A(w_t)}{\alpha}$ *except when* w_t *is below a threshold, which is negative if* $z < 1$.

(iii) *An increase in* $\frac{u^\top \Sigma u}{\rho}$ *raises* $A(w_t)$.

Proof: We first prove Part (i). Differentiating (3.20) with respect to α, and noting that α also affects z, we find

$$\frac{\partial A(w_t)}{\partial \alpha} = \frac{2\sqrt{z}\cot\left(\frac{\alpha w_t}{\sqrt{z}}\right) - 1}{1 + z} - \frac{2z\left(\sqrt{z}\cot\left(\frac{\alpha w_t}{\sqrt{z}}\right) - 1\right)}{(1+z)^2}$$

$$= \frac{2\sqrt{z}\cot\left(\frac{\alpha w_t}{\sqrt{z}}\right) + z - 1}{(1+z)^2}. \tag{A.49}$$

(All partial derivatives with respect to α in this and subsequent proofs take into account the dependence of z on α, instead of treating z as a constant.) Since $\cot(x)$ is decreasing for $x \in (0, \frac{\pi}{2})$, the numerator in (A.49) is larger than

$$2\sqrt{z}\cot\left(\frac{\alpha \bar{w}}{\sqrt{z}}\right) + z - 1 = 1 + z > 0,$$

where the first step follows from (3.21). Therefore, an increase in α raises $A(w_t)$.

We next prove Part (ii). Using (3.20), we find

$$\frac{\partial \frac{A(w_t)}{\alpha}}{\partial \alpha} = \frac{\sqrt{z}(1 - z)\cot\left(\frac{\alpha w_t}{\sqrt{z}}\right) + 2z}{\alpha(1+z)^2}. \tag{A.50}$$

The numerator in (A.50) is positive for $z < 1$. For $z > 1$, the numerator is increasing in w_t, converges to $-\infty$ when w_t goes to zero, and is equal to $1 + z > 0$ for $w_t = \bar{w}$ because of (3.21).

47

Therefore, for $z > 1$, the numerator is negative for w_t below a threshold $\bar{w}_2 < \bar{w}$ and is positive for $w_t > \bar{w}_2$. The effect of α on $\frac{A(w_t)}{\alpha}$ is thus as in the lemma.

We finally prove Part (iii). Differentiating (3.20) with respect to $\frac{u^\top \Sigma u}{\rho}$ is equivalent to differentiating with respect to z holding α constant. We compute the derivative with respect to $\sqrt{z}$, which has the same sign as that with respect to z. We find

$$\frac{\partial A(w_t)}{\partial \sqrt{z}} = \alpha \frac{(1-z)\cot\left(\frac{\alpha w_t}{\sqrt{z}}\right) + 2\sqrt{z} + \frac{\alpha(1+z)w_t}{\sqrt{z}}\left(1 + \cot^2\left(\frac{\alpha w_t}{\sqrt{z}}\right)\right)}{(1+z)^2}. \tag{A.51}$$

We set $\hat{w} \equiv \frac{\alpha w_t}{\sqrt{z}}$ and write the numerator in (A.51) as

$$N(\hat{w}, z) \equiv (1-z)\cot(\hat{w}) + 2\sqrt{z} + (1+z)\hat{w}\left(1 + \cot^2(\hat{w})\right).$$

The function $N(\hat{w}, z)$ is positive for $z \leq 1$. Therefore, it is positive for all $z > 0$ if its derivative with respect to z

$$\frac{\partial N(\hat{w}, z)}{\partial z} = -\cot(\hat{w}) + \frac{1}{\sqrt{z}} + \hat{w}\left(1 + \cot^2(\hat{w})\right)$$

is positive. Eq. (A.16) implies that for $\hat{w}$ close to zero,

$$\frac{\partial N(\hat{w}, z)}{\partial z} = -\frac{1}{\hat{w}}\left(1 - \frac{\hat{w}^2}{3}\right) + \frac{1}{\sqrt{z}} + \hat{w}\left(1 + \frac{1}{\hat{w}^2}\left(1 - \frac{\hat{w}^2}{3}\right)^2\right) + o(\hat{w}) = \frac{1}{\sqrt{z}} + o(1) > 0.$$

Therefore, $\frac{\partial N(\hat{w}, z)}{\partial z}$ is positive if its derivative with respect to $\hat{w}$

$$\frac{\partial^2 N(\hat{w}, z)}{\partial \hat{w} \partial z} = 2\left(1 + \cot^2(\hat{w})\right) - 2\hat{w}\cot(\hat{w})\left(1 + \cot^2(\hat{w})\right) = 2\left(1 - \hat{w}\cot(\hat{w})\right)\left(1 + \cot^2(\hat{w})\right)$$

is positive. Since the function $x\cot(x)$ is decreasing, it is smaller than one, its limit when x goes to zero (Lemma A.1). Therefore, $\frac{\partial^2 N(\hat{w}, z)}{\partial \hat{w} \partial z}$ is positive, and so is $N(\hat{w}, z)$, implying that an increase in z raises $A(w_t)$.

Proof of Proposition 3.8: The results for the logarithmic case follow from (3.14), (3.15), and $A(w_t) = \frac{1}{w_t}$. We next prove the results for the limit risk-neutral case. The result for the Sharpe ratio in Part (i) follows from (3.15) and because an increase in α raises $\frac{\alpha A(w_t)}{\alpha + A(w_t)}$. The latter property

follows from

$$\frac{\partial}{\partial \alpha} \left(\frac{\alpha A(w_t)}{\alpha + A(w_t)} \right) = \frac{A(w_t)^2 + \alpha^2 \frac{\partial A(w_t)}{\partial \alpha}}{(\alpha + A(w_t))^2}$$

and Part (i) of Lemma A.2. The result for arbitrageur positions in Part (i) follows from (3.14) and Part (ii) of Lemma A.2. The result for the Sharpe ratio in Part (ii) follows from (3.15) and Part (iii) of Lemma A.2. The result for arbitrageur positions in Part (ii) follows from (3.14) and Part (iii) of Lemma A.2. ∎

Proof of Lemma 4.1: Using (2.1) and (4.1), we can write (4.4) and (4.5) as

$$dv_t = rv_t dt + X_t^\top (\mu_{St} + \bar{D} - rS_t) dt + \left(X_t^\top (\sigma_{St} + \sigma)^\top + u^\top \sigma^\top \right) dB_t, \tag{A.52}$$

$$dw_t = (rw_t - c_t) dt + Y_t^\top (\mu_{St} + \bar{D} - rS_t) dt + Y_t^\top (\sigma_{St} + \sigma)^\top dB_t, \tag{A.53}$$

respectively. If S_t, X_t, and Y_t satisfy (4.6), (4.7), and (4.8), then (A.52) is identical to (A.1), and (A.53) to (A.5). Therefore, if x_t and y_t maximize the objective of hedgers and of arbitrageurs, respectively, given π_t, then the same is true for X_t and Y_t, given S_t. Moreover, if x_t and y_t satisfy the market-clearing equation (3.12), then X_t and Y_t satisfy the market-clearing equation

$$X_t + Y_t = 0 \tag{A.54}$$

because of (4.7) and (4.8). Since (A.52) is identical to (A.1), and (A.53) to (A.5), the dynamics of arbitrageur wealth and the exposures of hedgers and arbitrageurs to the Brownian shocks are the same in the equilibrium (S_t, X_t, Y_t) as in (π_t, x_t, y_t). The arbitrageurs' Sharpe ratio is also the same since its numerator is $y_t^\top (\bar{D} - \pi_t)$ with short-lived assets and $Y_t^\top (\mu_{St} + \bar{D} - rS_t)$ with long-lived assets, and its denominator is $\sqrt{y_t^\top \sigma^\top \sigma y_t}$ with short-lived assets and $\sqrt{Y_t^\top (\sigma_{St} + \sigma)^\top (\sigma_{St} + \sigma) Y_t}$ with long-lived assets. ∎

Proof of Proposition 4.1: Setting $S_t = S(w_t)$ and combining Ito's lemma with (4.1), we find

$$\mu_{St} = \mu_{wt} S'(w_t) + \frac{1}{2} \sigma_{wt}^\top \sigma_{wt} S''(w_t)$$

$$= \left(r - q(w_t)^{-\frac{1}{\gamma}} \right) w_t S'(w_t) + \frac{\alpha^2}{(\alpha + A(w_t))^2} u^\top \Sigma u \left(A(w_t) S'(w_t) + \frac{1}{2} S''(w_t) \right), \tag{A.55}$$

where the second step follows from (A.29), and

$$\sigma_{St} = \sigma_{wt} S'(w_t)^\top$$

$$= \frac{\alpha}{\alpha + A(w_t)} \sigma u S'(w_t)^\top, \tag{A.56}$$

where the second step follows from (A.30).

Multiplying (4.6) from the left by $(\sigma_{St} + \sigma)^\top$, and using (3.13), we find

$$\mu_{St} + \bar{D} - rS_t = \frac{\alpha A(w_t)}{\alpha + A(w_t)}(\sigma_{St} + \sigma)^\top \sigma u. \tag{A.57}$$

Substituting (A.55) and (A.56) into (A.57), we find the ODE

$$\left(r - q^{-\frac{1}{\gamma}}\right)wS' + \frac{\alpha^2}{2(\alpha + A)^2}u^\top \Sigma u S'' + \bar{D} - rS = \frac{\alpha A}{\alpha + A}\Sigma u. \tag{A.58}$$

Assuming that $S(w_t)$ is given by (4.9) and substituting into (A.58), we find that $g(w_t)$ solves the ODE (4.10).

Substituting μ_{St} from (A.55) into (4.2), and using (4.9) and (4.10), we can write the instantaneous expected return as (4.11). Substituting σ_{St} from (A.56) into (4.3), and using (4.9), we can write the instantaneous covariance matrix as (4.12). ∎

Proof of Proposition 4.2: We first compute $g(w_t)$ in the limit logarithmic case. Noting that $q(w_t) = \frac{1}{\rho}$ and $A(w_t) = \frac{1}{w_t}$, and taking the limit when r goes to zero, we can write (4.10) as

$$-\rho w g' + \frac{\alpha^2 w^2}{2(\alpha w + 1)^2}u^\top \Sigma u g'' = -\frac{\alpha^2 w}{\alpha w + 1}$$

$$\Leftrightarrow -\frac{(\alpha w + 1)^2}{zw}g' + g'' = -\frac{2(\alpha w + 1)}{u^\top \Sigma u w}. \tag{A.59}$$

Multiplying both sides of (A.59) by the integrating factor

$$\exp\left(-\int \frac{(\alpha w + 1)^2}{zw}dw\right) = w^{-\frac{1}{z}}\exp\left(-\frac{1}{2z}\left(\alpha^2 w^2 + 4\alpha w\right)\right),$$

we find

$$\left[g'w^{-\frac{1}{z}}\exp\left(-\frac{1}{2z}\left(\alpha^2 w^2 + 4\alpha w\right)\right)\right]' = -\frac{2(\alpha w + 1)}{u^\top \Sigma u w}w^{-\frac{1}{z}}\exp\left(-\frac{1}{2z}\left(\alpha^2 w^2 + 4\alpha w\right)\right). \tag{A.60}$$

Integrating (A.60) once with the boundary condition that $g'(w_t) = 0$ remains bounded when w_t goes to ∞, we find

$$g'(w_t)w_t^{-\frac{1}{z}}\exp\left(-\frac{1}{2z}\left(\alpha^2 w_t^2 + 4\alpha w_t\right)\right) = \frac{2}{u^\top \Sigma u}\int_{w_t}^\infty \left(\alpha + \frac{1}{w}\right)w^{-\frac{1}{z}}\exp\left(-\frac{1}{2z}\left(\alpha^2 w^2 + 4\alpha w\right)\right)dw,$$

50

which yields (4.14).

We next compute $g'(w_t)$ in the limit risk-neutral case. Noting that $q(w_t)^{-\frac{1}{\gamma}}$ is equal to zero in $(0, \bar{w})$, and taking the limit when r goes to zero, we can write (4.10) as

$$\frac{\alpha^2}{2(\alpha + A)^2} u^\top \Sigma u g'' = -\frac{\alpha^2}{\alpha + A}$$

$$\Leftrightarrow g'' = -\frac{2(\alpha + A)}{u^\top \Sigma u}. \tag{A.61}$$

Integrating (A.61) once with the boundary condition $g'(\bar{w}) = 0$, we find

$$g'(w_t) = \frac{2}{u^\top \Sigma u} \int_{w_t}^{\bar{w}} (\alpha + A(w)) dw. \tag{A.62}$$

The boundary condition is implied by smooth-pasting and because $g(w_t)$ is independent of w_t for $w_t \geq \bar{w}$. Using (3.20) to compute the integral in (A.62), we find (4.15). ∎

Proof of Proposition 4.3: We first show the results in the limit risk-neutral case. Eq. (4.11) implies that Part (i) holds if the function

$$K(w_t) \equiv \frac{\alpha A(w_t)}{\alpha + A(w_t)} \left[f(w_t) u^\top \Sigma u + 1 \right] \tag{A.63}$$

is increasing in w_t for $w_t < \bar{w}_a$ and decreasing for $w_t > \bar{w}_a$. The derivative of $K(w_t)$ with respect to w_t is

$$K'(w_t) = \frac{\alpha^2}{(\alpha + A(w_t))^2} \left\{ A'(w_t) \left[\frac{(\alpha - A(w_t)) g'(w_t)}{\alpha + A(w_t)} u^\top \Sigma u + 1 \right] + A(w_t) g''(w_t) u^\top \Sigma u \right\} \tag{A.64}$$

$$= \frac{\alpha^2}{(\alpha + A(w_t))^2} \left\{ A'(w_t) \left[\frac{(\alpha - A(w_t)) g'(w_t)}{\alpha + A(w_t)} u^\top \Sigma u + 1 \right] - 2A(w_t)(\alpha + A(w_t)) \right\}, \tag{A.65}$$

where (A.64) follows from (4.13), and (A.65) because (A.62) implies that

$$g''(w_t) = -\frac{2(\alpha + A(w_t))}{u^\top \Sigma u}. \tag{A.66}$$

Since $A'(w_t) < 0$, the term $A(w_t) - \alpha$ is positive when w_t is below the threshold $\bar{w}_c$ defined by $A(\bar{w}_c) = \alpha$ and is negative when $w > \bar{w}_c$. Therefore, (A.65) implies that $K'(w_t) < 0$ for $w_t \geq \bar{w}_c$.

For $w_t < \bar{w}_c$, $K'(w_t)$ has the same sign as

$$K_1(w_t) \equiv g'(w_t)u^\top \Sigma u - \frac{\alpha + A(w_t)}{A(w_t) - \alpha} + \frac{2A(w_t)(\alpha + A(w_t))^2}{A'(w_t)(A(w_t) - \alpha)}$$

$$= g'(w_t)u^\top \Sigma u - \frac{\alpha + A(w_t)}{A(w_t) - \alpha} - \frac{2A(w_t)(\alpha + A(w_t))^2}{\left(A(w_t)^2 + \frac{(\alpha + A(w_t))^2}{z}\right)(A(w_t) - \alpha)}, \tag{A.67}$$

where the second step follows from (A.20). The function $K_1(w_t)$ converges to ∞ when w_t goes to zero because $g'(w_t)$ and $A(w_t)$ converge to ∞, and converges to $-\infty$ when w_t goes to $\bar{w}_c$ from below. If, therefore, $K_1(w_t)$ is decreasing in w_t, it is positive when w_t is below a threshold $\bar{w}_a \in (0, \bar{w}_c)$ and is negative when $w_t > \bar{w}_a$. The first term in (A.67) is decreasing in w_t because (A.66) implies that $g'(w_t)$ is decreasing. The second term is increasing in w_t because $A(w_t)$ is decreasing in w_t and the function

$$x \to \frac{\alpha + x}{x - \alpha}$$

is decreasing in x for $x \in (\alpha, \infty)$. Likewise, the third term is increasing in w_t if the function

$$x \to \frac{2x(\alpha + x)^2}{\left(x^2 + \frac{(\alpha + x)^2}{z}\right)(x - \alpha)}$$

is decreasing in x for $x \in (\alpha, \infty)$. The derivative of the latter function with respect to x has the same sign as

$$\left[(\alpha + x)^2 + 2x(\alpha + x)\right]\left(x^2 + \frac{(\alpha + x)^2}{z}\right)(x - \alpha)$$

$$- x(\alpha + x)^2\left[2\left(x + \frac{\alpha + x}{z}\right)(x - \alpha) + \left(x^2 + \frac{(\alpha + x)^2}{z}\right)\right]$$

$$= \alpha(\alpha + x)\left[x^2(\alpha - 3x) - \frac{(\alpha + x)^3}{z}\right],$$

which is negative for $x \in (\alpha, \infty)$. Therefore, $K_1(w_t)$ is decreasing in w_t, and so $K(w_t)$ is increasing in w_t for $w_t < \bar{w}_a$ and decreasing for $w_t > \bar{w}_a$.

Eq. (4.12) implies that Parts (ii) and (iii) hold if $f(w_t)$ is increasing in w_t for $w_t < \bar{w}_b$ and decreasing for $w_t > \bar{w}_b$. If, in particular, such a threshold $\bar{w}_b$ exists, it is larger than the threshold $\bar{w}_a$ in Part (i) because $K(w_t)$ is the product of $\frac{\alpha A(w_t)}{\alpha + A(w_t)}$, which is decreasing in w_t, times $f(w_t)u^\top \Sigma u + 1$.

The derivative of $f(w_t)$ with respect to w_t is

$$f'(w_t) = \frac{\alpha}{(\alpha + A(w_t))^2} \left[-A'(w_t)g'(w_t) + g''(w_t)(\alpha + A(w_t)) \right] \tag{A.68}$$

$$= -\frac{\alpha}{(\alpha + A(w_t))^2} \left[A'(w_t)g'(w_t) + \frac{2(\alpha + A(w_t))^2}{u^\top \Sigma u} \right], \tag{A.69}$$

where (A.68) follows from the definition of $f(w_t)$, and (A.69) from (A.66). Since $A'(w_t) < 0$, (A.69) implies that $f'(w_t)$ has the same sign as

$$H_1(w_t) \equiv g'(w_t)u^\top \Sigma u + \frac{2(\alpha + A(w_t))^2}{A'(w_t)}$$

$$= g'(w_t)u^\top \Sigma u - \frac{2(\alpha + A(w_t))^2}{A(w_t)^2 + \frac{(\alpha + A(w_t))^2}{z}}$$

$$= g'(w_t)u^\top \Sigma u - \frac{2z}{z \left[\frac{A(w_t)}{\alpha + A(w_t)} \right]^2 + 1}, \tag{A.70}$$

where the second step follows from (A.20). The function $H_1(w_t)$ converges to ∞ when w_t goes to zero because $g'(w_t)$ and $A(w_t)$ converge to ∞, and converges to $-2z < 0$ when w_t goes to $\bar{w}$ because $g'(w_t)$ and $A(w_t)$ converge to zero. If, therefore, $H_1(w_t)$ is decreasing in w_t, it is positive when w_t is below a threshold $\bar{w}_b$ and is negative when $w_t > \bar{w}_b$. The first term in (A.70) is decreasing in w_t because $g'(w_t)$ is decreasing. The second term is increasing in w_t because $A(w_t)$ is decreasing. Therefore, $H_1(w_t)$ is decreasing in w_t, and so $f(w_t)$ is increasing in w_t for $w_t < \bar{w}_b$ and decreasing for $w_t > \bar{w}_b$.

To show Part (iv), we use (4.12) to write the correlation as

$$\text{Corr}_t(dR_{nt}, dR_{n't}) = \frac{f(w_t) \left[f(w_t)u^\top \Sigma u + 2 \right] (\Sigma u)_n (\Sigma u)_{n'} + \Sigma_{nn'}}{\sqrt{\left\{ f(w_t) \left[f(w_t)u^\top \Sigma u + 2 \right] (\Sigma u)_n^2 + \Sigma_{nn} \right\} \left\{ f(w_t) \left[f(w_t)u^\top \Sigma u + 2 \right] (\Sigma u)_{n'}^2 + \Sigma_{n'n'} \right\}}}. \tag{A.71}$$

Differentiating (A.71) with respect to $f(w_t)$, we find that $\text{Corr}_t(dR_{nt}, dR_{n't})$ is increasing in $f(w_t)$ if (4.16) holds and is decreasing in $f(w_t)$ if (4.16) holds in the opposite direction. Part (iv) then follows from the behavior of $f(w_t)$ shown in the proof of Parts (ii) and (iii).

To show Part (v), we use (4.9) and (A.56) to write (4.8) as

$$\sigma y_t = \sigma \left(I + f(w_t) u u^\top \Sigma \right) Y_t$$

$$\Leftrightarrow y_t = \left(I + f(w_t) u u^\top \Sigma \right) Y_t$$

$$\Leftrightarrow Y_t + f(w_t) u^\top \Sigma Y_t u = \frac{\alpha}{\alpha + A(w_t)} u, \tag{A.72}$$

where I is the $N \times N$ identity matrix, and the third step follows from (3.14). Eq. (A.72) implies that Y_t is collinear with u. Setting $Y_t = \nu u$ in (A.72), we find

$$\frac{\alpha}{\alpha + A(w_t)} = \nu + f(w_t) \nu u^\top \Sigma u \Rightarrow \nu = \frac{\alpha}{(\alpha + A(w_t))(1 + f(w_t) u^\top \Sigma u)},$$

and so

$$Y_t = \frac{\alpha}{(\alpha + A(w_t))(1 + f(w_t) u^\top \Sigma u)} u = \frac{\alpha}{\alpha + A(w_t) + \alpha g'(w_t) u^\top \Sigma u} u. \tag{A.73}$$

Part (v) follows from (A.73) and because $A(w_t)$ and $g'(w_t)$ are decreasing in w_t.

We next show the results in the limit logarithmic case. We start by determining the asymptotic behavior of $g'(w_t)$ for w_t close to zero and w_t close to ∞. For w close to zero, the integrand in (4.14) is

$$w^{-1-\frac{1}{z}} + o \left(w^{-1-\frac{1}{z}} \right).$$

Hence, for w_t close to zero, the integral in (4.14) is

$$z w_t^{-\frac{1}{z}} + o \left(w_t^{-\frac{1}{z}} \right),$$

and (4.14) implies that

$$\lim_{w_t \to 0} g'(w_t) = \frac{2z}{u^\top \Sigma u}. \tag{A.74}$$

To determine the asymptotic behavior for w_t close to ∞, we set $w = w_t + x$ and write the integral

54

in (4.14) as

$$
\int_0^\infty \left(\alpha + \frac{1}{w_t + x} \right) (w_t + x)^{-\frac{1}{z}} \exp\left(-\frac{1}{2z} \left(\alpha^2 (w_t + x)^2 + 4\alpha(w_t + x) \right) \right) dx
$$

$$
= w_t^{-\frac{1}{z}} \exp\left(-\frac{1}{2z} \left(\alpha^2 w_t^2 + 4\alpha w_t \right) \right)
$$

$$
\times \int_0^\infty \left(\alpha + \frac{1}{w_t + x} \right) \left(1 + \frac{x}{w_t} \right)^{-\frac{1}{z}} \exp\left(-\frac{1}{2z} \left(2\alpha^2 w_t x + \alpha^2 x^2 + 4\alpha x \right) \right) dx. \tag{A.75}
$$

We can further write the integral in (A.75) as

$$
\int_0^\infty Q\left(x, \frac{1}{w_t} \right) \exp(-Rw_t x) dx, \tag{A.76}
$$

where

$$
Q(x, y) \equiv \left(\alpha + \frac{y}{1 + xy} \right) (1 + xy)^{-\frac{1}{z}} \exp\left(-\frac{1}{2z} \left(\alpha^2 x^2 + 4\alpha x \right) \right),
$$

$$
R \equiv \frac{\alpha^2}{z}.
$$

Because of the term $\exp(-Rw_t x)$, the behavior of the integral (A.76) for large w_t is determined by the behavior of the function $Q(x, y)$ for (x, y) close to zero. We set

$$
Q(x, y) = Q(0, 0) + \frac{\partial Q}{\partial x}(0, 0)x + \frac{\partial Q}{\partial y}(0, 0)y + \hat{Q}(x, y), \tag{A.77}
$$

where $\hat{Q}(x, y)$ involves terms of order two and higher in (x, y). Substituting (A.77) into (A.76), and integrating, we find

$$
\int_0^\infty Q\left(x, \frac{1}{w_t} \right) \exp(-Rw_t x) dx
$$

$$
= Q(0, 0)\frac{1}{Rw_t} + \frac{\partial Q}{\partial x}(0, 0)\frac{1}{R^2 w_t^2} + \frac{\partial Q}{\partial y}(0, 0)\frac{1}{Rw_t^2} + \int_0^\infty \hat{Q}\left(x, \frac{1}{w_t} \right) \exp(-Rw_t x) dx. \tag{A.78}
$$

Since

$$
Q(0, 0) = \alpha,
$$

$$
\frac{\partial Q}{\partial x}(0, 0) = -\frac{2\alpha^2}{z},
$$

$$
\frac{\partial Q}{\partial y}(0, 0) = 1,
$$

and the integral in $\hat{Q}$ yields terms of order smaller than $\frac{1}{w_t^2}$ for large w_t, (A.78) implies that

$$\int_0^\infty Q\left(x, \frac{1}{w_t}\right) \exp(-Rw_t x) dx = \frac{z}{\alpha w_t} - \frac{z}{\alpha^2 w_t^2} + o\left(\frac{1}{w_t^2}\right). \tag{A.79}$$

Substituting back into (A.75) and then back into (4.14), we find that for w_t close to ∞,

$$g'(w_t) = \frac{2z}{u^\top \Sigma u \alpha w_t} - \frac{2z}{u^\top \Sigma u \alpha^2 w_t^2} + o\left(\frac{1}{w_t^2}\right). \tag{A.80}$$

We next show that $g'(w_t)$ is decreasing in w_t. Assume, by contradiction, that there exists $\underline{w}$ such that $g''(\underline{w}) \geq 0$. Since $g'(w_t)$ is positive and converges to zero when w_t converges to ∞, there exists $\overline{w} > \underline{w}$ such that $g''(\overline{w}) < 0$. Therefore, the function $g''(w_t)$ must cross the x-axis from above in $[\underline{w}, \overline{w})$, i.e., there must exist $\hat{w} \in [\underline{w}, \overline{w})$ such that $g''(\hat{w}) = 0$ and $g'''(\hat{w}) \leq 0$. Since $g'(w_t)$ satisfies the ODE (A.59), it also satisfies

$$-\frac{\alpha}{z}g' - \frac{\alpha w + 1}{z}g'' + \frac{d}{dw}\left(\frac{w}{\alpha w + 1}\right)g'' + \frac{w}{\alpha w + 1}g''' = 0, \tag{A.81}$$

which follows from (A.59) by multiplying both sides by $\frac{w}{\alpha w + 1}$ and differentiating with respect to w. Eq. (A.81) cannot hold at $\hat{w}$ because $g'(\hat{w}) > 0$, $g''(\hat{w}) = 0$, and $g'''(\hat{w}) \leq 0$, a contradiction. Therefore, $g''(w_t) < 0$ for all w_t.

Part (v) follows from the arguments in the limit risk-neutral case and because the functions $A(w_t) = \frac{1}{w_t}$ and $g'(w_t)$ are positive and decreasing in w_t. Part (i) also follows from the arguments in that case if the function $K(w_t)$ defined by (A.63) is decreasing in w_t in the case $z < \frac{1}{2}$, and is increasing in w_t for $w_t < \bar{w}_a$ and decreasing for $w_t > \bar{w}_a$ in the case $z > \frac{1}{2}$. Using $A(w_t) = \frac{1}{w_t}$, we can write the derivative of $K(w_t)$ with respect to w_t, given by (A.64), as

$$K'(w_t) = \frac{\alpha^2}{(\alpha w_t + 1)^2}\left[\frac{(1 - \alpha w_t)g'(w_t)}{\alpha w_t + 1}u^\top \Sigma u - 1 + w_t g''(w_t)u^\top \Sigma u\right] \tag{A.82}$$

$$= \frac{\alpha^2}{(\alpha w_t + 1)^3}\left[\left(1 - \alpha w_t + \frac{(\alpha w_t + 1)^3}{z}\right)g'(w_t)u^\top \Sigma u - (2\alpha w_t + 3)(\alpha w_t + 1)\right], \tag{A.83}$$

where the second step follows by substituting $g''(w_t)$ from (A.59). Eqs. (A.82) and $g'(w_t) < 0$ imply

that $K'(w_t) < 0$ for $w_t \geq \frac{1}{\alpha}$. For $w_t < \frac{1}{\alpha}$, (4.14) and (A.83) imply that $K'(w_t)$ has the same sign as

$$K_2(w_t) \equiv 2 \int_{w_t}^{\infty} \left(\alpha + \frac{1}{w} \right) w^{-\frac{1}{z}} \exp \left(-\frac{1}{2z} \left(\alpha^2 w^2 + 4\alpha w \right) \right) dw$$

$$- \frac{(2\alpha w_t + 3)(\alpha w_t + 1) w_t^{-\frac{1}{z}} \exp \left(-\frac{1}{2z} \left(\alpha^2 w_t^2 + 4\alpha w_t \right) \right)}{1 - \alpha w_t + \frac{(\alpha w_t + 1)^3}{z}}.$$

The derivative of $K_2(w_t)$ with respect to w_t is

$$K_2'(w_t) = -2 \left(\alpha + \frac{1}{w_t} \right) w_t^{-\frac{1}{z}} \exp \left(-\frac{1}{2z} \left(\alpha^2 w_t^2 + 4\alpha w_t \right) \right) - \frac{w_t^{-\frac{1}{z}} \exp \left(-\frac{1}{2z} \left(\alpha^2 w_t^2 + 4\alpha w_t \right) \right)}{\left(1 - \alpha w_t + \frac{(\alpha w_t + 1)^3}{z} \right)^2}$$

$$\times \left[\left(\alpha(4\alpha w_t + 5) - \frac{(2\alpha w_t + 3)(\alpha w_t + 1)^3}{z w_t} \right) \left(1 - \alpha w_t + \frac{(\alpha w_t + 1)^3}{z} \right) \right.$$

$$\left. -\alpha \left(-1 + \frac{3(\alpha w_t + 1)^2}{z} \right) (2\alpha w_t + 3)(\alpha w_t + 1) \right],$$

and has the same sign as

$$K_3(w_t) \equiv -\frac{2(\alpha^2 w_t^2 + 3\alpha w_t + 1)}{(\alpha w_t + 1)^3} + \frac{4\alpha^2 w_t^2 + 3\alpha w_t - 1}{z} + \frac{(\alpha w_t + 1)^3}{z^2}.$$

The function $K_3(w_t)$ is equal to

$$K_3(0) = -2 - \frac{1}{z} + \frac{1}{z^2} = \frac{(1 - 2z)(1 + z)}{z^2}$$

for $w_t = 0$, and is increasing in w_t because the function $\frac{\alpha^2 w_t^2 + 3\alpha w_t + 1}{(\alpha w_t + 1)^3}$ is decreasing in w_t. The function $K_2(w_t)$ is equal to

$$K_2(w_t) = 2z w_t^{-\frac{1}{z}} - \frac{3z}{1 + z} w_t^{-\frac{1}{z}} + o \left(w_t^{-\frac{1}{z}} \right) = \frac{(2z - 1)z}{1 + z} w_t^{-\frac{1}{z}} + o \left(w_t^{-\frac{1}{z}} \right)$$

for w_t close to zero. Moreover, $K_2 \left(\frac{1}{\alpha} \right) < 0$ because the functions $K'(w_t)$ and $K_2(w_t)$ have the same sign for $w_t \leq \frac{1}{\alpha}$ and (A.82) implies that $K' \left(\frac{1}{\alpha} \right) < 0$.

- When $z < \frac{1}{2}$, $K_3(0) > 0$ and $K_3(w_t)$ increasing in w_t imply that $K_3(w_t) > 0$. Therefore, $K_2(w_t)$ is increasing in w_t. Since $K_2 \left(\frac{1}{\alpha} \right) < 0$, $K_2(w_t)$ is negative for $w_t < \frac{1}{\alpha}$. Therefore, $K(w_t)$ is decreasing in w_t for all $w_t \in (0, \infty)$.

57

- When $z > \frac{1}{2}$, $K_3(0) < 0$ and $K_3(w_t)$ increasing in w_t imply that $K_3(w_t) < 0$ for $w_t \in (0, \frac{1}{\alpha})$ except possibly in an interval ending at $\frac{1}{\alpha}$. Therefore, $K_2(w_t)$ is decreasing in w_t for $w_t \in (0, \frac{1}{\alpha})$ except possibly in an interval ending at $\frac{1}{\alpha}$ where it is increasing. Since $K_2(w_t)$ is positive for w_t close to zero and $K_2\left(\frac{1}{\alpha}\right) < 0$, $K_2(w_t)$ is positive when w_t is below a threshold $\bar{w}_a \in (0, \frac{1}{\alpha})$ and negative when $w_t \in (\bar{w}_a, \frac{1}{\alpha})$. Therefore, $K(w_t)$ is increasing in w_t for $w_t \in (0, \bar{w}_a)$ and decreasing for $w_t \in (\bar{w}_a, \infty)$.

Parts (ii), (iii), and (iv) follow from the arguments in the limit risk-neutral case if the function $f(w_t)$ is increasing in w_t for $w_t < \bar{w}_b$ and decreasing for $w_t > \bar{w}_b$. Using $A(w_t) = \frac{1}{w_t}$ and substituting $g''(w_t)$ from (A.59), we can write the derivative of $f(w_t)$ with respect to w_t, given by (A.68), as

$$f'(w_t) = \frac{\alpha}{(\alpha w_t + 1)^2}\left[\left(\frac{(\alpha w_t + 1)^3}{z} + 1\right)g'(w_t) - \frac{2(\alpha w_t + 1)^2}{u^\top \Sigma u}\right]. \tag{A.84}$$

Eqs. (4.14) and (A.84) imply that $f'(w_t)$ has the same sign as

$$H_2(w_t) \equiv \int_{w_t}^{\infty}\left(\alpha + \frac{1}{w}\right)w^{-\frac{1}{z}}\exp\left(-\frac{1}{2z}\left(\alpha^2 w^2 + 4\alpha w\right)\right)dw - \frac{(\alpha w_t + 1)^2 w_t^{-\frac{1}{z}}\exp\left(-\frac{1}{2z}\left(\alpha^2 w_t^2 + 4\alpha w_t\right)\right)}{\frac{(\alpha w_t + 1)^3}{z} + 1}.$$

The derivative of $H_2(w_t)$ with respect to w_t is

$$H_2'(w_t) = -\left(\alpha + \frac{1}{w_t}\right)w_t^{-\frac{1}{z}}\exp\left(-\frac{1}{2z}\left(\alpha^2 w_t^2 + 4\alpha w_t\right)\right) - \frac{w_t^{-\frac{1}{z}}\exp\left(-\frac{1}{2z}\left(\alpha^2 w_t^2 + 4\alpha w_t\right)\right)}{\left(\frac{(\alpha w_t + 1)^3}{z} + 1\right)^2}$$

$$\times \left[\left(2\alpha(\alpha w_t + 1) - \frac{(\alpha w_t + 1)^4}{z w_t}\right)\left(\frac{(\alpha w_t + 1)^3}{z} + 1\right) - \frac{3\alpha(\alpha w_t + 1)^4}{z}\right],$$

and has the same sign as

$$H_3(w_t) \equiv -\frac{2\alpha w_t + 1}{(\alpha w_t + 1)^3} + \frac{\alpha w_t - 1}{z}.$$

The function $H_3(w_t)$ is negative for $w_t = 0$ and converges to ∞ when w_t goes to ∞. Moreover, it is increasing in w_t because the function $\frac{2\alpha w_t + 1}{(\alpha w_t + 1)^3}$ is decreasing in w_t. Therefore, $H_3(w_t) < 0$ when w_t is below a threshold $\bar{w}_d$ and $H_3(w_t) > 0$ when $w_t > \bar{w}_d$. For w_t close to zero, the function $H_2(w_t)$ is equal to

$$H_2(w_t) = z w_t^{-\frac{1}{z}} - \frac{z}{1 + z}w_t^{-\frac{1}{z}} + o\left(w_t^{-\frac{1}{z}}\right) = \frac{z^2}{1 + z}w_t^{-\frac{1}{z}} + o\left(w_t^{-\frac{1}{z}}\right),$$

58

and hence is positive. Moreover, $H_2(w_t)$ converges to zero when w_t goes to ∞. Since $H_2(w_t)$ is decreasing in w_t for $w_t < \bar{w}_d$ and increasing for $w_t > \bar{w}_d$, it is positive when w_t is below a threshold $\bar{w}_b < \bar{w}_d$ and negative when $w_t > \bar{w}_b$. Therefore, $f(w_t)$ is increasing in w_t for $w_t < \bar{w}_b$ and decreasing for $w_t > \bar{w}_b$. ∎

Proof of Proposition 5.1: Using (4.9), (A.54), and (A.73) to compute the partial derivatives in (5.1), we find (5.2). In the limit when r goes to zero, λ_{nt} converges to

$$
\left(1 + \frac{A(w_t)}{\alpha} + g'(w_t) u^\top \Sigma u \right) \alpha \Sigma_{nn}.
$$

This expression is decreasing in w_t because $A(w_t)$ is decreasing (shown in the proof of Proposition 3.4 for the limit risk-neutral case) and $g(w_t)$ is decreasing (shown in the proof of Proposition 4.3).

Proof of Corollary 5.1: We set

$$
\lambda_{nt} = \left(1 + \frac{A(w_t)}{\alpha} + g'(w_t) u^\top \Sigma u \right) (\alpha - rg(w_t)) \Sigma_{nn} \equiv L(w_t) \Sigma_{nn}. \tag{A.85}
$$

Using (A.85) and Ito's lemma, we find

$$
\mathrm{Cov}_t(d\Lambda_t, dR_t) = L'(w_t) \frac{\alpha \sum_{n'=1}^{N} \Sigma_{n'n'}}{N} \mathrm{Cov}_t(dw_t, dR_t), \tag{A.86}
$$

$$
\mathrm{Cov}_t(d\Lambda_t, d\lambda_{nt}) = \left(L'(w_t) \right)^2 \frac{\alpha^2 \Sigma_{nn} \sum_{n'=1}^{N} \Sigma_{n'n'}}{N} \mathrm{Var}_t(dw_t), \tag{A.87}
$$

$$
\mathrm{Cov}_t(d(u^\top dR_t), d\lambda_{nt}) = L'(w_t) \alpha \Sigma_{nn} u^\top \mathrm{Cov}_t(dw_t, dR_t). \tag{A.88}
$$

The diffusion matrix of the return vector dR_t is

$$
(\sigma_{St} + \sigma)^\top = \left(\frac{\alpha}{\alpha + A(w_t)} \sigma u S'(w_t)^\top + \sigma \right)^\top
$$

$$
= \left(\frac{\alpha g'(w_t)}{\alpha + A(w_t)} \sigma u u^\top \Sigma + \sigma \right)^\top, \tag{A.89}
$$

where the first step follows from (A.56) and the second from (4.9). The covariance between wealth and the return vector dR_t is

$$
\mathrm{Cov}_t(dw_t, dR_t) = (\sigma_{St} + \sigma)^\top \sigma_{wt}
$$

$$
= \left(\frac{\alpha g'(w_t)}{\alpha + A(w_t)} \sigma u u^\top \Sigma + \sigma \right)^\top \frac{\alpha}{\alpha + A(w_t)} \sigma u
$$

$$
= \frac{\alpha}{\alpha + A(w_t)} \left[f(w_t) u^\top \Sigma u + 1 \right] \Sigma u, \tag{A.90}
$$

where the second step follows from (A.30) and (A.89). Part (i) of the corollary follows by substituting (A.90) into (A.86). The proportionality coefficient is

$$C^\Lambda(w_t) = L'(w_t) \frac{\alpha^2 \sum_{n'=1}^{N} \Sigma_{n'n'}}{N(\alpha + A(w_t))} \left[f(w_t) u^\top \Sigma u + 1 \right], \tag{A.91}$$

and is negative in the limit when r goes to zero because $L(w_t)$ is decreasing in w_t. Part (ii) of the corollary follows from (A.87). The proportionality coefficient is positive for any r. Part (i) of the corollary follows by substituting (A.90) into (A.88). The proportionality coefficient is negative in the limit when r goes to zero because $L(w_t)$ is decreasing in w_t. ∎

Proof of Corollary 5.2: The proportionality result follows from (4.11), (A.86), and (A.90). These equations imply that the proportionality coefficient is

$$\Pi^\Lambda(w_t) = \frac{A(w_t)}{L'(w_t) \frac{\alpha \sum_{n'=1}^{N} \Sigma_{n'n'}}{N}}. \tag{A.92}$$

This coefficient is negative in the limit when r goes to zero because $L(w_t)$ is decreasing in w_t. ∎

Proof of Proposition 5.2: In the limit when r goes to zero, (A.85) implies that $L(w_t)$ converges to

$$\left(1 + \frac{A(w_t)}{\alpha} + g'(w_t) u^\top \Sigma u \right) \alpha.$$

Substituting into (A.91) and (A.92), we find

$$C^\Lambda(w_t) = \left(\frac{A'(w_t)}{\alpha} + g''(w_t) u^\top \Sigma u \right) \frac{\alpha^3 \sum_{n'=1}^{N} \Sigma_{n'n'}}{N(\alpha + A(w_t))} \left[f(w_t) u^\top \Sigma u + 1 \right], \tag{A.93}$$

$$\Pi^\Lambda(w_t) = \frac{A(w_t)}{\left(\frac{A'(w_t)}{\alpha} + g''(w_t) u^\top \Sigma u \right) \frac{\alpha^2 \sum_{n'=1}^{N} \Sigma_{n'n'}}{N}}. \tag{A.94}$$

We first show the properties of $C^\Lambda(w_t)$ and $\Pi^\Lambda(w_t)$ in the limit risk-neutral case. Using (4.13), (A.20), (A.61), and (A.62), we can write (A.93) and (A.94) as

$$C^\Lambda(w_t) = - \left[\frac{A(w_t)^2 + \frac{(\alpha + A(w_t))^2}{z}}{\alpha} + 2(\alpha + A(w_t)) \right] \frac{\alpha^3 \sum_{n'=1}^{N} \Sigma_{n'n'}}{N(\alpha + A(w_t))} \left[\frac{2\alpha \int_{w_t}^{\bar{w}} (\alpha + A(w)) dw}{\alpha + A(w_t)} + 1 \right], \tag{A.95}$$

$$\Pi^\Lambda(w_t) = - \frac{A(w_t)}{\left[\frac{A(w_t)^2 + \frac{(\alpha + A(w_t))^2}{z}}{\alpha} + 2(\alpha + A(w_t)) \right] \frac{\alpha^2 \sum_{n'=1}^{N} \Sigma_{n'n'}}{N}}. \tag{A.96}$$

When w_t goes to zero, $A(w_t)$ converges to ∞. Therefore, (A.95) implies that $C^\Lambda(w_t)$ converges to $-\infty$, and (A.96) implies that $\Pi^\Lambda(w_t)$ converges to zero. For $w_t = \bar{w}$, $A(\bar{w}) = 0$. Therefore, (A.95) implies that

$$C^\Lambda(\bar{w}) = -\left(\frac{1}{z} + 2\right)\frac{\alpha^3 \sum_{n'=1}^N \Sigma_{n'n'}}{N} < 0$$

and (A.96) implies that $\Pi^\Lambda(\bar{w}) = 0$. To show the inverse hump shape of $\Pi^\Lambda(w_t)$, we write (A.96) as

$$\Pi^\Lambda(w_t) = -\frac{1}{\left[\frac{A(w_t)}{\alpha}\left(1 + \frac{1}{z}\right) + 2\left(1 + \frac{1}{z}\right) + \frac{\alpha}{A(w_t)}\left(2 + \frac{1}{z}\right)\right]\frac{\alpha^2 \sum_{n'=1}^N \Sigma_{n'n'}}{N}}. \tag{A.97}$$

The term in square brackets in the denominator of (A.97) is an inverse hump-shaped function of $A(w_t)$. Since $A(w_t)$ is decreasing in w_t, $\Pi^\Lambda(w_t)$ is an inverse hump-shaped function of w_t.

We next show the properties of $C^\Lambda(w_t)$ and $\Pi^\Lambda(w_t)$ in the limit logarithmic case. Using $A(w_t) = \frac{1}{w_t}$, (4.13), and (A.59), we can write (A.93) and (A.94) as

$$C^\Lambda(w_t) = \left[-\frac{1}{\alpha w_t^2} + \frac{\alpha w_t + 1}{w_t}\left(\frac{(\alpha w_t + 1)u^\top \Sigma u}{z}g'(w_t) - 2\right)\right]\frac{\alpha^3 \sum_{n'=1}^N \Sigma_{n'n'}}{N\left(\alpha + \frac{1}{w_t}\right)}\left[\frac{\alpha g'(w_t)}{\alpha + \frac{1}{w_t}}u^\top \Sigma u + 1\right], \tag{A.98}$$

$$\Pi^\Lambda(w_t) = \frac{\frac{1}{w_t}}{\left[-\frac{1}{\alpha w_t^2} + \frac{\alpha w_t + 1}{w_t}\left(\frac{(\alpha w_t + 1)u^\top \Sigma u}{z}g'(w_t) - 2\right)\right]\frac{\alpha^2 \sum_{n'=1}^N \Sigma_{n'n'}}{N}}. \tag{A.99}$$

When w_t goes to zero, $g'(w_t)$ converges to the positive limit (A.74). Therefore, (A.98) implies that $C^\Lambda(w_t)$ converges to $-\infty$, and (A.99) implies that $\Pi^\Lambda(w_t)$ converges to zero. When w_t goes to ∞, (A.80) implies that $g'(w_t)$ is of order $\frac{1}{w_t}$ and

$$\frac{(\alpha w_t + 1)u^\top \Sigma u}{z}g'(w_t) - 2 = -\frac{2}{\alpha^2 w_t^2} + o\left(\frac{1}{w_t^2}\right).$$

Therefore, (A.98) implies that $C^\Lambda(w_t)$ converges to zero, and (A.99) implies that $\Pi^\Lambda(w_t)$ converges to $-\infty$. ∎

Proof of Proposition 6.1 First we show that $q(w_t), q_1(w_t)$ and $g(w_t)$ must solve the system of

ODEs

$$\rho q = \gamma q^{1-\frac{1}{\gamma}} + \left(r - q^{-\frac{1}{\gamma}}\right) q'w + rq(1-\gamma) + \tag{A.100}$$

$$+\frac{1}{2}\left(q'' + \frac{2q'\gamma}{w} - \frac{2q'^2}{q} + \frac{q(1-\gamma)\gamma}{w^2}\right) \frac{\alpha^2}{\left(A\left(w_t\right) + \alpha - \alpha g'\left(w\right)\left(u+s\right)^\top \Sigma^\top s\right)^2} \left(u+s\right)^\top \Sigma \left(u+s\right)$$

or, for $\gamma = 1$,

$$\rho q_1 = \log(\rho) + \frac{r-\rho}{\rho} + (r-\rho)q_1' + \frac{1}{2}\left(q_1'' + \frac{2q_1'}{w} + \frac{1}{\rho w^2}\right) \frac{\alpha^2}{\left(\alpha + \frac{1}{w} - \alpha g'\left(w\right)\left(u+s\right)^\top \Sigma^\top s\right)^2} \left(u+s\right)^\top \Sigma \left(u+s\right)$$

$$\tag{A.101}$$

and

$$\frac{\alpha A\left(w_t\right)}{\alpha + A\left(w_t\right) - g'\left(w_t\right)\left(u+s\right)^\top \Sigma^\top s\alpha} =$$

$$= \left(r - q(w_t)^{-\frac{1}{\gamma}}\right) w_t g'\left(w\right) + \frac{\alpha^2}{\alpha + A(w_t)} \frac{\left(u+s\right)^\top \Sigma \left(u+s\right)}{\alpha + A(w_t) - g'\left(w_t\right)\left(u+s\right)^\top \Sigma^\top s\alpha} \frac{1}{2} g''\left(w_t\right) + \alpha - rg\left(w_t\right)$$

where we set $q(w_t)^{-\frac{1}{\gamma}} = \rho$ for $\gamma = 1$ and the definition of $A\left(w_t\right)$ and the boundary conditions do not change and that in equilibrium

$$c_t = q(w_t)^{-\frac{1}{\gamma}} w_t$$

$$\left(\sigma_{S_t} + \sigma\right) X_t + \sigma u = \frac{A\left(w_t\right)}{A\left(w_t\right) + \alpha} b_t \tag{A.102}$$

$$\left(\sigma_{S_t} + \sigma\right) Y_t = \frac{\alpha}{A\left(w_t\right) + r\alpha} b_t \tag{A.103}$$

where

$$b_t = \frac{\left(A\left(w_t\right) + \alpha\right)\sigma\left(u+s\right)}{A\left(w_t\right) + \alpha - \alpha g'\left(w\right)\left(u+s\right)^\top \Sigma^\top s} \tag{A.104}$$

and

$$\sigma_{S_t}^\top = \frac{\alpha g'\left(w_t\right)}{A\left(w_t\right) + \alpha - \alpha g'\left(w\right)\left(u+s\right)^\top \Sigma^\top s} \sigma\left(u+s\right)\left(u+s\right)^\top \Sigma^\top \tag{A.105}$$

62

and

$$f\left(w_t\right) = \frac{\alpha g'\left(w_t\right)}{A\left(w_t\right) + \alpha - \alpha g'\left(w\right)\left(u+s\right)^\top \Sigma^\top s}.$$

Given that the dynamics of hedgers' and arbitrageurs' wealth follows an Ito process where the drifts and diffusions are

$$\mu_{vt} = rv_t - c_{ht} + X_t^\top\left(\mu_{S_t} + \bar{D} - rS_t\right) + u^\top \bar{D}, \tag{A.106}$$

$$\sigma_{vt} = \left(X_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right) + u^\top\sigma^\top\right),$$

for the hedger and

$$\mu_{wt} = \left(rw_t - c_t + Y_t^\top\left(\mu_{S_t} + \bar{D} - rS_t\right)\right), \tag{A.107}$$

$$\sigma_{wt} = Y_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right),$$

for arbitrageurs,.the first-order condition for the hedger gives

$$\left(\sigma_{S_t} + \sigma\right) X_t + \sigma u = \frac{A\left(w_t\right)}{A\left(w_t\right) + \alpha}b_t \tag{A.108}$$

where

$$b_t \equiv \left(\sigma_{S_t} + \sigma\right) s + \sigma u.$$

Turning to arbitrageurs, the Bellman equation is

$$\rho V = \max_{c_t, y_t}\left\{u(\hat{c}_t) + V_{\hat{w}}\mu_{\hat{w}t} + \frac{1}{2}V_{\hat{w}\hat{w}}\sigma_{\hat{w}t}^\top\sigma_{\hat{w}t} + V_w\mu_{wt} + \frac{1}{2}V_{ww}\sigma_{wt}^\top\sigma_{wt} + V_{\hat{w}w}\sigma_{\hat{w}t}^\top\sigma_{wt}\right\}, \tag{A.109}$$

where $u(\hat{c}_t) = \frac{\hat{c}_t^{1-\gamma}}{1-\gamma}$ for $\gamma \neq 1$ and $u(\hat{c}_t) = \log(\hat{c}_t)$ for $\gamma = 1$, $(\mu_{\hat{w}t}, \sigma_{\hat{w}t})$ are the drift and diffusion of the arbitrageur's own wealth $\hat{w}_t$, and (μ_{wt}, σ_{wt}) are the drift and diffusion of the arbitrageurs' total wealth.

When $\gamma \neq 1$, we substitute (3.6) and (A.106)-(A.9) into (A.109) to write it as

$$\rho q(w_t)\frac{\hat{w}_t^{1-\gamma}}{1-\gamma} = \max_{\hat{c}_t, \hat{y}_t}\left\{\begin{array}{c}\frac{\hat{c}_t^{1-\gamma}}{1-\gamma} + q(w_t)\hat{w}_t^{-\gamma}\left(r\hat{w}_t - \hat{c}_t + \hat{Y}_t^\top\left(\mu_{S_t} + \bar{D} - rS_t\right)\right) \\ -\frac{1}{2}q(w_t)\gamma\hat{w}_t^{-\gamma-1}\hat{Y}_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right)\left(\sigma_{S_t} + \sigma\right)\hat{Y}_t \\ +q'(w_t)\frac{\hat{w}_t^{1-\gamma}}{1-\gamma}\left(rw_t - c_t + Y_t^\top\left(\mu_{S_t} + \bar{D} - rS_t\right)\right) \\ +\frac{1}{2}q''(w_t)\frac{\hat{w}_t^{1-\gamma}}{1-\gamma}Y_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right)\left(\sigma_{S_t} + \sigma\right)Y_t + q'(w_t)\hat{w}_t^{-\gamma}\hat{Y}_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right)\left(\sigma_{S_t} + \sigma\right)Y_t\end{array}\right\}$$

The first-order conditions with respect to $\hat{c}_t$ and $\hat{Y}_t$ are

$$\hat{c}_t = q(w_t)^{-\frac{1}{\gamma}}\hat{w}_t \tag{A.111}$$

$$\hat{Y}_t = \frac{\hat{w}_t}{\gamma}\left(\left(\left(\sigma_{S_t}^\top + \sigma^\top\right)\left(\sigma_{S_t} + \sigma\right)\right)^{-1}\left(\mu_{S_t} + \bar{D} - rS_t\right) + \frac{q'(w_t)}{q(w_t)}Y_t\right)$$

, respectively. When $\gamma = 1$, we substitute (3.7),(A.106)-(A.9) into (A.109) to write it as

$$\rho\left(\frac{1}{\rho}\log(\hat{w}_t) + q_1(w_t)\right) = \max_{\hat{c}_t,\hat{y}_t}\left\{\begin{array}{c}\log(\hat{c}_t) + \frac{1}{\rho\hat{w}_t}\left(r\hat{w}_t - \hat{c}_t + \hat{Y}_t^\top(\mu_{S_t} + \bar{D} - rS_t)\right)\\ -\frac{1}{2\rho\hat{w}_t^2}\hat{Y}_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right)\left(\sigma_{S_t} + \sigma\right)\hat{Y}_t + q_1'(w_t)\left(rw_t - c_t + Y_t^\top(\mu_{S_t} + \bar{D} - rS_t)\right)\\ +\frac{1}{2}q_1''(w_t)Y_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right)\left(\sigma_{S_t} + \sigma\right)Y_t\end{array}\right\}$$

The first-order conditions with respect to $\hat{c}_t$ and $\hat{Y}_{tt}$ are (A.111) for $q(w_t) = \frac{1}{\rho}$. Since in equilibrium $\hat{c}_t = c_t$ and $\hat{w}_t = w_t$, (A.111) implies that

$$c_t = q(w_t)^{-\frac{1}{\gamma}}w_t, \tag{A.112}$$

and using the definition of $A(w_t) = \frac{\gamma}{w_t} - \frac{q'(w_t)}{q(w_t)}$, we find

$$\left(\sigma_{S_t}^\top + \sigma^\top\right)Y_t = \frac{\left(((\sigma_{S_t} + \sigma))^{-1}(\mu_{S_t} + \bar{D} - rS_t)\right)}{A(w_t)}. \tag{A.113}$$

Then, (A.108), (A.113) and the market clearing equation gives

$$\left(\sigma_{S_t}^\top + \sigma^\top\right)^{-1}\left(\mu_{S_t} + \bar{D} - rS_t\right) = \frac{A(w_t)\alpha}{A(w_t) + \alpha}b_t \tag{A.114}$$

implying also

$$Y_t^\top\left(\mu_{S_t} + \bar{D} - rS_t\right) = \frac{A(w_t)\alpha^2}{(A(w_t) + \alpha)^2}b_t^\top b_t \tag{A.115}$$

$$X_t^\top\left(\mu_{S_t} + \bar{D} - rS_t\right) = \left(\frac{A(w_t)\alpha(A(w_t))}{(A(w_t) + \alpha)^2}b_t^\top b_t - \frac{A(w_t)\alpha}{A(w_t) + \alpha}u^\top\sigma^\top b_t\right) \tag{A.116}$$

Substituting (A.114)-(A.116) and (A.112) and the definition of $A(w_t)$ into (A.110) and simplifying gives

$$\rho q = \gamma q^{1-\frac{1}{\gamma}} + \left(r - q^{-\frac{1}{\gamma}}\right) q'w + rq(1-\gamma) + \frac{1}{2}\left(q'' + \frac{2q'\gamma}{w} - \frac{2q'^2}{q} + \frac{q(1-\gamma)\gamma}{w^2}\right) \frac{\alpha^2}{(A(w_t)+\alpha)^2} b_t^\top b_t.$$

(A.117)

Before, we simplify (A.117) further, we derive the ODE for $g(w_t)$. Note that from Ito's Lemma

$$\mu_{St} = \mu_{wt} S'(w_t) + \frac{1}{2}\sigma_{wt}^\top \sigma_{wt} S''(w_t) = (rw_t - q(w_t)^{-\frac{1}{\gamma}} w_t + \frac{A(w_t)\alpha^2}{(A(w_t)+\alpha)^2} b_t^\top b_t) S'(w_t)$$

$$+ \frac{1}{2}(r\alpha)^2 \frac{1}{(A(w_t)+\alpha)^2} b_t^\top b_t S''(w_t) =$$

$$= \left(r - q(w_t)^{-\frac{1}{\gamma}}\right) w_t S'(w_t) + \frac{\alpha^2}{(A(w_t)+\alpha)^2} b_t^\top b_t \left(A(w_t)S'(w_t) + \frac{1}{2}S''(w_t)\right),$$

and

$$\sigma_{St} = \sigma_{wt}^\top S'(w_t)^\top = (\sigma_{S_t} + \sigma) Y_t S'(w_t)^\top$$

(A.118)

$$= \frac{\alpha}{A(w_t)+\alpha} b_t S'(w_t)^\top,$$

(A.119)

Substituting back σ_{S_t} into the definition of b_t, we get

$$b_t = \frac{\sigma(u+s)}{\left(1 - \frac{\alpha}{A(w_t)+\alpha}S'(w_t)^\top s\right)}.$$

(A.120)

and using (4.9), we get (A.104) and (A.105). Substituting (A.104) and (A.105) into (A.114), we find the ODE for $g(w_t)$ (A.100). Also, given (A.104), and (A.117) simplifies to (A.100). Using the same steps for the $\gamma = 1$ case, we get (A.101). Finally, using (A.104) and (A.105) and the definition of $\frac{E_t(dR_t)}{dt}$ and (A.114), we get (6.2). ∎

Proof of Propositions 6.2 First we show that the functions $q(w_t), q_1(w_t), g(w_t), F(w_t)$ must

solve the system of ODEs

$$\rho q \;=\; \gamma q^{1-\frac{1}{\gamma}} + \left(r - q^{-\frac{1}{\gamma}}\right) q' w + rq(1-\gamma) + \tag{A.121}$$

$$+\frac{1}{2}\frac{\left(q'' + \frac{2q'\gamma}{w} - \frac{2q'^2}{q} + \frac{q(1-\gamma)\gamma}{w^2}\right)(r\alpha)^2}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha - r\alpha g'\left(w\right)\left(u+s\right)^\top \Sigma^\top s\right)^2}\left(u+s\right)^\top \Sigma\left(u+s\right)$$

or, for $\gamma = 1$,

$$\rho q_1 = \log(\rho) + \frac{r-\rho}{\rho} + (r-\rho)q_1' + \frac{1}{2}\frac{\left(q_1'' + \frac{2q_1'}{w} + \frac{1}{\rho w^2}\right)(r\alpha)^2}{\left(\frac{1}{w} - F'\left(w_t\right) + r\alpha - r\alpha g'\left(w\right)\left(u+s\right)^\top \Sigma^\top s\right)^2}\left(u+s\right)^\top \Sigma\left(u+s\right)$$

$$\tag{A.122}$$

and

$$\frac{A\left(w_t\right)r\alpha}{A\left(w_t\right) - F'\left(w_t\right) + r\alpha - r\alpha g'\left(w\right)\left(u+s\right)^\top \Sigma^\top s} = \tag{A.123}$$

$$= \left(r - q(w_t)^{-\frac{1}{\gamma}}\right) w_t g'(w_t) + \frac{(r\alpha)^2\left(u+s\right)^\top \Sigma\left(u+s\right)}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha - r\alpha g'\left(w\right)\left(u+s\right)^\top \Sigma^\top s\right)^2}\frac{1}{2}g''(w_t)$$

$$+r\alpha - rg\left(w_t\right).$$

and

$$0 \;=\; \rho - r - rF(w_t) + r\log(r) + rau^\top \bar{D} + F'(w_t)(r - q(w_t)^{-\frac{1}{\gamma}})w_t + \tag{A.124}$$

$$+\frac{(r\alpha)^2\frac{1}{2}\left(F'' + A\left(w_t\right)^2\right)}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha - r\alpha g'\left(w\right)\left(u+s\right)^\top \Sigma^\top s\right)^2}\left(u+s\right)^\top \Sigma\left(u+s\right)$$

$$-\frac{A\left(w_t\right)(r\alpha)^2}{A\left(w_t\right) - F'\left(w_t\right) + r\alpha - r\alpha g'\left(w\right)\left(u+s\right)^\top \Sigma^\top s}u^\top \Sigma\left(u+s\right) \tag{A.125}$$

and that

$$c_t = q(w_t)^{-\frac{1}{\gamma}}w_t$$

$$\left(\sigma_{S_t} + \sigma\right) X_t + \sigma u = \frac{A\left(w_t\right) - F'\left(w_t\right)}{A\left(w_t\right) - F'\left(w_t\right) + r\alpha} b_t \tag{A.126}$$

$$\left(\sigma_{S_t} + \sigma\right) Y_t = \frac{r\alpha}{A\left(w_t\right) - F'\left(w_t\right) + r\alpha} b_t \tag{A.127}$$

where

$$b_t = \frac{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha\right)\sigma\left(u + s\right)}{A\left(w_t\right) - F'\left(w_t\right) + r\alpha - r\alpha g'\left(w\right)\left(u + s\right)^\top \Sigma^\top s}. \tag{A.128}$$

and

$$\sigma_{S_t} = \frac{r\alpha g'\left(w_t\right)}{A\left(w_t\right) + r\alpha - F'\left(w_t\right) - r\alpha g'\left(w\right)\left(u + s\right)^\top \Sigma^\top s}\sigma\left(u + s\right)\left(u + s\right)^\top \Sigma^\top \tag{A.129}$$

and

$$f\left(w_t\right) = \frac{r\alpha g'\left(w_t\right)}{A\left(w_t\right) - F'\left(w_t\right) + r\alpha - r\alpha g'\left(w\right)\left(u + s\right)^\top \Sigma^\top s}.$$

The Bellman equation of a hedger is

$$\rho_h V^h = \max_{c_{ht}, x_t}\left\{ u(c_{ht}) + V_v^h \mu_{vt} + \frac{1}{2}V_{vv}^h \sigma_{vt}^\top \sigma_{vt} + V_w^h \mu_{wt} + \frac{1}{2}V_{ww}^h \sigma_{wt}^\top \sigma_{wt} + V_{vw}^h \sigma_{vt}^\top \sigma_{wt} \right\}, \tag{A.130}$$

where $u(c_{ht}) = e^{-\alpha c_{ht}}$, (μ_{vt}, σ_{vt}) are the drift and diffusion of the hedger's wealth v_t, and (μ_{wt}, σ_{wt}) are the drift and diffusion of arbitrageur wealth given by (A.8) and (A.9).

We substitute (6.3) and (A.8)-(A.9) into (A.130) to write it as

$$-\rho e^{-[r\alpha v_t + F(w_t)]} = \max_{c_{ht}, x_t}\left\{ \begin{array}{c} -e^{-\alpha c_{hs}} + r\alpha e^{-[r\alpha v_t + F(w_t)]}\left(rv_t - c_{ht} + X_t^\top\left(\mu_{S_t} + \bar{D} - rS_t\right) + u^\top \bar{D}\right) \\ -\frac{1}{2}(r\alpha)^2 e^{-[r\alpha v_t + F(w_t)]}\left(\left(X_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right) + u^\top \sigma^\top\right)\left(X_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right) + u^\top \sigma^\top\right)^\top \\ +F'(w_t)e^{-[r\alpha v_t + F(w_t)]}\left(rw_t - c_t + Y_t^\top\left(\mu_{S_t} + \bar{D} - rS_t\right)\right) \\ +\frac{1}{2}\left[F''(w_t) - F'(w_t)^2\right]e^{-[r\alpha v_t + F(w_t)]}\left(Y_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right)\right)\left(\sigma_{S_t}^\top + \sigma^\top\right)^\top Y_t \\ -r\alpha F'(w_t)e^{-[r\alpha v_t + F(w_t)]}\left(X_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right) + u^\top \sigma^\top\right)\left(\sigma_{S_t}^\top + \sigma^\top\right)^\top Y_t \end{array} \right\}$$

$$\tag{A.131}$$

The first-order conditions with respect to c_{ht} and X_t are

$$e^{-\alpha c_{hs}} = re^{-[r\alpha v_t + F(w_t)]}, \tag{A.132}$$

$$\mu_{S_t} + \bar{D} - rS_t = r\alpha\left(\sigma_{S_t}^\top + \sigma^\top\right)\left(X_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right) + u^\top \sigma^\top\right)^\top + F'(w_t)\left(\sigma_{S_t}^\top + \sigma^\top\right)\left(\sigma_{S_t}^\top + \sigma^\top\right)^\top Y_t, \tag{A.133}$$

respectively. Which we can rewrite as

$$c_{hs} = rv_t + \frac{F(w_t)}{\alpha} - \frac{\log(r)}{\alpha}, \tag{A.134}$$

$$\left(\sigma_{S_t} + \sigma\right) X_t + \sigma u = \frac{\left(\sigma_{S_t}^\top + \sigma^\top\right)^{-1} \left(\mu_{S_t} + \bar{D} - rS_t\right)}{r\alpha} - \frac{F'(w_t)\left(\sigma_{S_t}^\top + \sigma^\top\right)^\top Y_t}{r\alpha} \tag{A.135}$$

respectively. Using (A.134) and (A.135), we can simplify (A.131) to

$$0 = \rho - r - rF(w_t) + r\log(r) + r\alpha X_t^\top(\mu_{S_t} + \bar{D} - rS_t) + r\alpha u^\top \bar{D}$$

$$-\frac{1}{2}(r\alpha)^2\left(\left(X_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right) + u^\top\sigma^\top\right)\left((\sigma_{S_t} + \sigma)X_t + \sigma u\right) + F'(w_t)(rw_t - c_t + Y_t^\top(\mu_{S_t} + \bar{D} - rS_t))\right)$$

$$+\frac{1}{2}\left[F''(w_t) - F'(w_t)^2\right]\left(Y_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right)\right)\left(\sigma_{S_t}^\top + \sigma^\top\right)^\top Y_t$$

$$- r\alpha F'(w_t)\left(X_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right) + u^\top\sigma^\top\right)\left(\sigma_{S_t}^\top + \sigma^\top\right)^\top Y_t.$$

Turning to arbitrageurs, the Bellman equation is

$$\rho V = \max_{c_t, y_t} \left\{ u(\hat{c}_t) + V_{\hat{w}}\mu_{\hat{w}t} + \frac{1}{2}V_{\hat{w}\hat{w}}\sigma_{\hat{w}t}^\top\sigma_{\hat{w}t} + V_w\mu_{wt} + \frac{1}{2}V_{ww}\sigma_{wt}^\top\sigma_{wt} + V_{\hat{w}w}\sigma_{\hat{w}t}^\top\sigma_{wt} \right\}, \tag{A.136}$$

where $u(\hat{c}_t) = \frac{\hat{c}_t^{1-\gamma}}{1-\gamma}$ for $\gamma \neq 1$ and $u(\hat{c}_t) = \log(\hat{c}_t)$ for $\gamma = 1$, $(\mu_{\hat{w}t}, \sigma_{\hat{w}t})$ are the drift and diffusion of the arbitrageur's own wealth $\hat{w}_t$, and (μ_{wt}, σ_{wt}) are the drift and diffusion of the arbitrageurs' total wealth.

When $\gamma \neq 1$, we substitute (3.6) and (A.8)-(A.9) into (A.136) to write it as

$$\rho q(w_t)\frac{\hat{w}_t^{1-\gamma}}{1-\gamma} = \max_{\hat{c}_t, \hat{y}_t} \left\{ \begin{array}{c} \frac{\hat{c}_t^{1-\gamma}}{1-\gamma} + q(w_t)\hat{w}_t^{-\gamma}\left(r\hat{w}_t - \hat{c}_t + \hat{Y}_t^\top(\mu_{S_t} + \bar{D} - rS_t)\right) \\ -\frac{1}{2}q(w_t)\gamma\hat{w}_t^{-\gamma-1}\hat{Y}_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right)(\sigma_{S_t} + \sigma)\hat{Y}_t + q'(w_t)\frac{\hat{w}_t^{1-\gamma}}{1-\gamma}\left(rw_t - c_t + Y_t^\top(\mu_{S_t} + \bar{D} - rS_t)\right) \\ +\frac{1}{2}q''(w_t)\frac{\hat{w}_t^{1-\gamma}}{1-\gamma}Y_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right)(\sigma_{S_t} + \sigma)Y_t + q'(w_t)\hat{w}_t^{-\gamma}\hat{Y}_t^\top\left(\sigma_{S_t}^\top + \sigma^\top\right)(\sigma_{S_t} + \sigma)Y_t \end{array} \right\} \tag{A.137}$$

The first-order conditions with respect to $\hat{c}_t$ and $\hat{Y}_t$ are

$$\hat{c}_t = q(w_t)^{-\frac{1}{\gamma}}\hat{w}_t \tag{A.138}$$

$$\hat{Y}_t = \frac{\hat{w}_t}{\gamma}\left(\left(\left(\sigma_{S_t}^\top + \sigma^\top\right)(\sigma_{S_t} + \sigma)\right)^{-1}(\mu_{S_t} + \bar{D} - rS_t) + \frac{q'(w_t)}{q(w_t)}Y_t\right)$$

68

, respectively.

When $\gamma = 1$, we substitute (3.7),(A.8)-(A.9) into (A.136) to write it as

$$\rho \left(\frac{1}{\rho} \log(\hat{w}_t) + q_1(w_t) \right) = \max_{\hat{c}_t, \hat{y}_t} \left\{ \begin{array}{c} \log(\hat{c}_t) + \frac{1}{\rho \hat{w}_t} \left(r\hat{w}_t - \hat{c}_t + \hat{Y}_t^\top (\mu_{S_t} + \bar{D} - rS_t) \right) \\ -\frac{1}{2\rho \hat{w}_t^2} \hat{Y}_t^\top \left(\sigma_{S_t}^\top + \sigma^\top \right) (\sigma_{S_t} + \sigma) \hat{Y}_t + q_1'(w_t) \left(rw_t - c_t + Y_t^\top (\mu_{S_t} + \bar{D} - rS_t) \right) \\ + \frac{1}{2} q_1''(w_t) Y_t^\top \left(\sigma_{S_t}^\top + \sigma^\top \right) (\sigma_{S_t} + \sigma) Y_t \end{array} \right\}$$

(A.139)

The first-order conditions with respect to $\hat{c}_t$ and $\hat{Y}_{tt}$ are (A.138) for $q(w_t) = \frac{1}{\rho}$. Since in equilibrium $\hat{c}_t = c_t$ and $\hat{w}_t = w_t$, (A.138) implies that

$$c_t = q(w_t)^{-\frac{1}{\gamma}} w_t, \tag{A.140}$$

and using the definition of $A(w_t) = \frac{\gamma}{w_t} - \frac{q'(w_t)}{q(w_t)}$, we find

$$\left(\sigma_{S_t}^\top + \sigma^\top \right) Y_t = \frac{\left(((\sigma_{S_t} + \sigma))^{-1} (\mu_{S_t} + \bar{D} - rS_t) \right)}{A(w_t)}. \tag{A.141}$$

Then, (A.135) and (A.141) and the market clearing equation

$$\frac{\left(\sigma_{S_t}^\top + \sigma^\top \right)^{-1} (\mu_{S_t} + \bar{D} - rS_t)}{r\alpha} - \frac{F'(w_t) \left(\sigma_{S_t}^\top + \sigma^\top \right)^\top Y_t}{r\alpha} - \sigma u + \frac{\left((\sigma_{S_t} + \sigma)^\top \right)^{-1} (\mu_{S_t} + \bar{D} - rS_t)}{A(w_t)} = (\sigma_{S_t} + \sigma) s,$$

(A.142)

form a linear system of three equations in three unknowns of $\left(\sigma_{S_t}^\top + \sigma^\top \right)^{-1} (\mu_{S_t} + \bar{D} - rS_t)$, $(\sigma_{S_t} + \sigma) X_t + \sigma u$, $\left(\sigma_{S_t}^\top + \sigma^\top \right)^\top Y_t$. The solution to the system is (A.126), (A.127) and

$$\left(\sigma_{S_t}^\top + \sigma^\top \right)^{-1} (\mu_{S_t} + \bar{D} - rS_t) = \frac{A(w_t) r\alpha}{A(w_t) - F'(w_t) + r\alpha} b_t \tag{A.143}$$

and where

$$b_t \equiv (\sigma_{S_t} + \sigma) s + \sigma u$$

69

implying also

$$Y_t^\top \left(\mu_{S_t} + \bar{D} - rS_t \right) = \frac{A\left(w_t\right)\left(r\alpha\right)^2}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha\right)^2} b_t^\top b_t$$

$$X_t^\top \left(\mu_{S_t} + \bar{D} - rS_t \right) = \left(\frac{A\left(w_t\right) r\alpha \left(A\left(w_t\right) - F'\left(w_t\right)\right)}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha\right)^2} b_t^\top b_t - \frac{A\left(w_t\right) r\alpha}{A\left(w_t\right) - F'\left(w_t\right) + r\alpha} u^\top \sigma^\top b_t \right)$$

Using (A.143),(A.127) and (A.140) we can rewrite (A.137) as

$$0 = \rho - r - rF(w_t) + r\log(r) +$$

$$+ r\alpha \left(\frac{A\left(w_t\right) r\alpha \left(A\left(w_t\right) - F'\left(w_t\right)\right)}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha\right)^2} b_t^\top b_t - \frac{A\left(w_t\right) r\alpha}{A\left(w_t\right) - F'\left(w_t\right) + r\alpha} u^\top \sigma^\top b_t \right) + u^\top \bar{D}$$

$$- \frac{1}{2}(r\alpha)^2 \frac{\left(A\left(w_t\right) - F'\left(w_t\right)\right)^2}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha\right)^2} b_t^\top b_t + F'\left(w_t\right)\left(w_t\right)(rw_t - q(w_t)^{-\frac{1}{\gamma}})w_t$$

$$+ \frac{A\left(w_t\right)\left(r\alpha\right)^2}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha\right)^2} b_t^\top b_t + \frac{1}{2}\left[F''(w_t) - F'(w_t)^2\right] \frac{(r\alpha)^2}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha\right)^2} b_t^\top b_t$$

$$- r\alpha F'(w_t) \frac{r\alpha \left(A\left(w_t\right) - F'\left(w_t\right)\right)}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha\right)^2} b_t^\top b_t.$$

which simplifies to

$$0 = \rho - r - rF(w_t) + r\log(r) + rau^\top \bar{D} + F'(w_t)(r - q(w_t)^{-\frac{1}{\gamma}})w_t \tag{A.144}$$

$$+ \frac{(r\alpha)^2 \frac{1}{2}\left(F'' + A\left(w_t\right)^2\right)}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha\right)^2} b_t^\top b_t - \frac{A\left(w_t\right)\left(r\alpha\right)^2}{A\left(w_t\right) - F'\left(w_t\right) + r\alpha} u^\top \sigma^\top b_t.$$

Similarly, substituting (A.143), (A.127) and (A.140) and the definition of $A\left(w_t\right)$ into (A.137) and simplifying gives

$$\rho q = \gamma q^{1-\frac{1}{\gamma}} + \left(r - q^{-\frac{1}{\gamma}}\right) q'w + rq(1-\gamma) + \frac{1}{2}\left(q'' + \frac{2q'\gamma}{w} - \frac{2q'^2}{q} + \frac{q(1-\gamma)\gamma}{w^2} \right) \frac{(r\alpha)^2}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha\right)^2} b_t^\top b_t..$$

$$\tag{A.145}$$

Before, we simplify (A.144)-(A.145) further, we derive the ODE for $g(w_t)$. Note that from Ito's Lemma

$$\mu_{St} = \mu_{wt}S'(w_t) + \frac{1}{2}\sigma_{wt}^\top\sigma_{wt}S''(w_t) =$$

$$= (rw_t - q(w_t)^{-\frac{1}{\gamma}}w_t + \frac{A(w_t)(r\alpha)^2}{(A(w_t) - F'(w_t) + r\alpha)^2}b_t^\top b_t)S'(w_t) + \frac{1}{2}(r\alpha)^2\frac{1}{(A(w_t) - F'(w_t) + r\alpha)^2}b_t^\top b_t S''(w_t) =$$

$$= \left(r - q(w_t)^{-\frac{1}{\gamma}}\right)w_t S'(w_t) + \frac{(r\alpha)^2}{(A(w_t) - F'(w_t) + r\alpha)^2}b_t^\top b_t\left(A(w_t)S'(w_t) + \frac{1}{2}S''(w_t)\right),$$

and

$$\sigma_{St} = \sigma_{wt}^\top S'(w_t)^\top = (\sigma_{S_t} + \sigma)Y_t S'(w_t)^\top \tag{A.146}$$

$$= \frac{r\alpha}{A(w_t) - F'(w_t) + r\alpha}b_t S'(w_t)^\top, \tag{A.147}$$

Substituting back σ_{S_t} into the definition of b_t, we get

$$b_t = ((\sigma_{S_t} + \sigma)s + \sigma u) = \frac{r\alpha}{A(w_t) - F'(w_t) + r\alpha}b_t S'(w_t)^\top s + \sigma(u + s)$$

$$b_t = \frac{\sigma(u + s)}{\left(1 - \frac{r\alpha}{A(w_t) - F'(w_t) + r\alpha}S'(w_t)^\top s\right)}. \tag{A.148}$$

and using (4.9), we get (A.128) and (A.129). Substituting (A.128) and (A.129) into (A.143), we find the ODE for $g(w_t)$ (A.123). Also, given (A.128), (A.144) and (A.145) simplifies to (A.124) and (A.121). Using the same steps for the $\gamma = 1$ case, we get (A.122). Finally, using (A.128) and (A.129) and the definition of $\frac{E_t(dR_t)}{dt}$ and (A.143), we get (6.5). ∎

Proof of Proposition 6.3: We show the argument for the case with CARA objective. The argument for only positive supply follows analogously.

First, we show that

$$\lambda_{nt} = \left(1 + \frac{A(w_t) - F'(w_t)}{r\alpha} + g'(w)(u + s)^\top \Sigma u\right)(\alpha - g(w_t))\Sigma_{nn}. \tag{A.149}$$

71

From (4.9)

$$\frac{\partial S_{nt}}{\partial u_n} = -\left(\alpha - g(w_t)\right) \Sigma_{nn}.$$

For $\frac{\partial X_{nt}}{\partial u_n}$, we use (A.127),(A.143),(A.147) and (A.128) to rewrite

$$\left(\sigma_{S_t} + \sigma\right) Y_t = \frac{r\alpha}{A\left(w_t\right) - F'\left(w_t\right) + r\alpha} b_t$$

as

$$\sigma\left(r\alpha g'\left(w_t\right) \frac{\left(u + s\right)\left(u + s\right)^\top \Sigma}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha - r\alpha g'\left(w\right)\left(u + s\right)^\top \Sigma^\top s\right)} + I\right) Y_t =$$

$$= r\alpha \frac{\sigma\left(u + s\right)}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha - r\alpha g'\left(w\right)\left(u + s\right)^\top \Sigma^\top s\right)}$$

multiplying both sides by σ^{-1} shows that Y_t must be collinear with $\left(u + s\right)$. So let us write $Y_t = \left(u + s\right) v$ giving

$$r\alpha g'\left(w_t\right) \frac{\left(u + s\right)^\top \Sigma \left(u + s\right) v}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha - r\alpha g'\left(w\right)\left(u + s\right)^\top \Sigma^\top s\right)} + v =$$

$$= r\alpha \frac{1}{\left(A\left(w_t\right) - F'\left(w_t\right) + r\alpha - r\alpha g'\left(w\right)\left(u + s\right)^\top \Sigma^\top s\right)}$$

which implies

$$v = \frac{1}{g'\left(w_t\right)\left(u + s\right)^\top \Sigma u + \frac{A(w_t) - F'(w_t)}{r\alpha} + 1}.$$

Hence, by market clearing

$$\frac{\partial X_n}{\partial u_n} = -\frac{\partial Y_n}{\partial u_n} = \frac{1}{g'\left(w_t\right)\left(u + s\right)^\top \Sigma u + \frac{A(w_t) - F'(w_t)}{r\alpha} + 1}$$

72

giving the result.

Note from (6.5) that the expected return of asset n is proportional to the covariance $(\Sigma\,(u+s))_n$ between the asset and portfolio u. The covariance between asset returns and aggregate illiquidity $\Lambda_t \equiv \frac{\sum_{n=1}^{N} \lambda_{nt}}{N}$ is

$$\frac{\mathrm{Cov}_t(d\Lambda_t, dR_t)}{dt} = C^\Lambda(w_t)\Sigma\,(u+s)\,, \tag{A.150}$$

where $C^\Lambda(w_t)$ is a negative coefficient. Given that both the expected returns in (6.2) and this covariance are proportional to $\Sigma\,(u+s)\,$, we can decompose the expected return of any asset n into a premium for this liquidity factor, $\Pi^\Lambda(w_t)$ and the liquidity beta as

$$\frac{\mathrm{E}_t(dR_t)}{dt} \equiv \Pi^\Lambda(w_t)\frac{\mathrm{Cov}_t(d\Lambda_t, dR_t)}{dt} \tag{A.151}$$

as claimed. ∎

References

Acharya, Viral V., and Lasse Heje Pedersen, 2005, Asset pricing with liquidity risk, *Journal of Financial Economics* 77, 375–410.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

Basak, Suleyman, and Anna Pavlova, 2013, Asset prices and institutional investors, *American Economic Review* 103, 1728–1758.

Bhamra, Harjoat S., and Raman Uppal, 2009, The effect of introducing a non-redundant derivative on the volatility of stock-market returns when agents differ in risk aversion, *Review of Financial Studies* 22, 2303–2330.

Brunnermeier, Markus, and Yuliy Sannikov, forthcoming, A macroeconomic model with a financial sector, *American Economic Review.*

Brunnermeier, Markus K., and Lasse Heje Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201–2238.

Buyuksahin, Bahattin, and Michel Robe, 2010, Speculators, commodities and cross-market linkages, http://papers.ssrn.com.

Chabakauri, Georgy, 2013, Asset pricing with heterogeneous investors and portfolio constraints, *Review of Financial Studies* 26, 3104–3141.

Chan, Yeung Lewis, and Leonid Kogan, 2002, Catching up with the joneses: Heterogeneous preferences and the dynamics of asset prices, *Journal of Political Economy* 110, 1255–1285.

Chen, Hui, Scott Joslin, and Sophie Ni, 2013, Demand for crash insurance and stock returns, http://www.mit.edu/ huichen/.

Cheng, Ing-Haw, Andrei Kirilenko, and Wei Xiong, 2012, Convective risk flows in commodity futures markets, .

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2000, Commonality in liquidity, *Journal of Financial Economics* 56, 3–28.

Comerton-Forde, Carole, Terrence Hendershott, Charles Jones, Pamela Moulton, and Mark Seasholes, 2010, Time variation in liquidity: The role of market-maker inventories and revenues, *Journal of Finance* 65, 295–331.

Danielsson, Jon, Hyun Song Shin, and Jean-Pierre Zigrand, 2012, Balance sheet capacity and endogenous risk, .

Dumas, Bernard, 1989, Two-person dynamic equilibrium in the capital market, *Review of Financial Studies* 2, 157–88.

Franzoni, Francesco, Eric Nowak, and Ludovic Phalippou, 2012, Private equity performance and liquidity risk, *Journal of Finance* 67, 2341–2373.

Froot, Kenneth, and Paul O'Connell, 1999, The pricing of US catastrophe reinsurance, in Kenneth Froot, ed.: *The Financing of Catastrophe Risk* (University of Chicago Press).

Garleanu, Nicolae, and Stavros Panageas, 2012, Young, old, conservative and bold: The implications of heterogeneity and finite lives for asset pricing, http://faculty.haas.berkeley.edu/garleanu/OLG.pdf.

Garleanu, Nicolae, and Lasse Heje Pedersen, 2011, Margin-based asset pricing and deviations from the law of one price, *Review of Financial Studies* 24, 1980–2022.

Gromb, Denis, and Dimitri Vayanos, 2002, Equilibrium and welfare in markets with financially constrained arbitrageurs, *Journal of Financial Economics* 66, 361–407.

———, 2014, The dynamics of financially constrained arbitrage, .

Hameed, Allaudeen, Wenjin Kang, and S. Viswanathan, 2010, Stock Market Declines and Liquidity, *Journal of Finance* 65, 257–293.

Hamilton, James D, and J Cynthia Wu, 2011, Risk premia in crude oil futures prices, .

Hasbrouck, Joel, and Duane Seppi, 2001, Common factors in prices, order flows and liquidity, *Journal of Financial Economics* 59, 383–411.

He, Zhiguo, and Arvind Krishnamurthy, 2013, Intermediary asset pricing, *American Economic Review* 103, 1–43.

Holmstrom, Bengt, and Jean Tirole, 2001, LAPM: A liquidity-based asset pricing model, *Journal of Finance* 56, 1837–1867.

Hong, Harrison, and Motohiro Yogo, 2012, What does futures market interest tell us about the macroeconomy and asset prices?, *Journal of Financial Economics* 105, 473 – 490.

Huberman, Gur, and Dominika Halka, 2001, Systematic liquidity, *Journal of Financial Research* 24, 161–178.

Irwin, Scott H., and Dwight R. Sanders, 2010, The impact of index and swap funds on commodity futures markets: Preliminary results, .

Jylha, Petri, and Matti Suominen, 2011, Speculative capital and currency carry trades, *Journal of Financial Economics* 99, 60–75.

Kondor, Peter, and Dimitri Vayanos, 2014, Liquidity risk and the dynamics of arbitrage capital: Extensions and robustness, http://www.personal.ceu.hu/staff/Peter_Kondor/papers.php.

Kyle, Albert, 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–1336.

Kyle, Albert S., and Wei Xiong, 2001, Contagion as a wealth effect, *Journal of Finance* 56, 1401–1440.

Longstaff, Francis A., and Jiang Wang, 2013, Asset pricing and the credit market, http://www.anderson.ucla.edu/faculty/finance/faculty/longstaff.

Merton, Robert, 1971, Optimum consumption and portfolio rules in a continuous-time model, *Journal of Economic Theory* 3, 373–413.

———— , 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.

Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

Sadka, Ronnie, 2010, Liquidity risk and the cross-section of hedge fund returns, *Journal of Financial Economics* 98, 54–71.

Tang, Ke, and Wei Xiong, 2010, Index investment and financialization of commodities, .

Vayanos, Dimitri, and Jiang Wang, 2013, Market liquidity—theory and empirical evidence, in George Constantinides, Milton Harris, and Rene Stulz, ed.: *Handbook of the Economics of Finance* (North Holland, Amsterdam).

Wang, Jiang, 1996, The term structure of interest rates in a pure exchange economy with heterogeneous investors, *Journal of Financial Economics* 41, 75–110.

Xiong, Wei, 2001, Convergence trading with wealth effects: an amplification mechanism in financial markets, *Journal of Financial Economics* 62, 247–292.