Liquidity Risk and the Dynamics of Arbitrage Capital

Péter Kondor
London School of Economics and CEPR

Dimitri Vayanos
London School of Economics, CEPR and NBER

March 29, 2018*

Abstract

We develop a continuous-time model of liquidity provision, in which hedgers can trade multiple risky assets with arbitrageurs. Arbitrageurs have CRRA utility, while hedgers' asset demand is independent of wealth. An increase in hedgers' risk aversion can make arbitrageurs endogenously more risk-averse. Because arbitrageurs generate endogenous risk, an increase in their wealth or a reduction in their CRRA coefficient can raise risk premia despite Sharpe ratios declining. Arbitrageur wealth is a priced risk factor because assets held by arbitrageurs offer high expected returns but suffer the most when wealth drops. Aggregate illiquidity, which declines in wealth, captures that factor.

*We thank Edina Berlinger, Bruno Biais, Itamar Dreschler, Nicolae Garleanu, Pete Kyle, Stavros Panageas, Anna Pavlova, Jean-Charles Rochet, Wei Xiong, Hongjun Yan, an anonymous associate editor and two anonymous referees; seminar participants at Atlanta Fed, Bank of Canada, Cass, Columbia, Copenhagen, Duke, HKU, Imperial, LBS, LSE, Minneapolis Fed, Montreal, NUS, Ohio State, Tinbergen, UIUC, UNC, Wash U, Wharton, and Zurich; and participants at the AEA, AFA, Arne Ryde, Cambridge INET, CRETE, EIEF Market Microstructure, ESEM, ESSFM Gerzensee, EWFC, NBER Asset Pricing, QMUL, and SED conferences, for helpful comments. We are very grateful to Mihail Zervos for guidance and encouragement with the proofs on differential equations. Kondor acknowledges financial support from the European Research Council (Starting Grant #336585). Both authors acknowledge financial support from the Paul Woolley Centre at the LSE.

1 Introduction

Liquidity in financial markets is often provided by specialized agents, such as market makers, trading desks in investment banks, and hedge funds. Adverse shocks to the capital of these agents cause liquidity to decline and risk premia to increase. Conversely, movements in the prices of assets held by liquidity providers feed back into these agents' capital.[1]

In this paper we study the dynamics of liquidity providers' capital, the liquidity that these agents provide to other participants, and assets' risk premia. We build a framework with minimal frictions, in particular no asymmetric information or borrowing constraints. The capital of liquidity providers matters in our model only because of standard wealth effects. At the same time, we depart from most frictionless asset-pricing models by fixing the riskless rate and by suppressing wealth effects for agents other than the liquidity providers. These assumptions are sensible when focusing on shocks to the capital of liquidity providers in an asset class rather than in the entire asset universe.

Our combination of assumptions makes it possible to prove general analytical results on equilibrium prices and allocations. We characterize, in particular, how liquidity providers' risk-appetite, the endogenous risk that they generate, and the pricing of that risk, depend on liquidity demanders' characteristics and on liquidity providers' capital. We also show that the capital of liquidity providers is the single priced risk factor, and that liquidity aggregated over the assets that we consider captures that factor because it increases in capital. Our results thus suggest that a priced liquidity risk factor may arise even with minimal frictions.

We assume a continuous-time infinite-horizon economy. There is a riskless asset with an exogenous and constant return, and multiple risky assets whose prices are determined endogenously in equilibrium. There are two sets of competitive agents: hedgers, who receive a risky income flow and seek to reduce their risk by participating in financial markets, and arbitrageurs, who take the other side of the trades that hedgers initiate. Arbitrageurs can be interpreted, for example, as speculators in futures markets. We consider two specifications for hedgers' preferences. Hedgers

[1] A growing empirical literature documents the relationship between the capital of liquidity providers, the liquidity that these agents provide to other participants, and assets' risk premia. 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. Aragon and Strahan (2012) find that following the collapse of Lehman Brothers in 2008, hedge funds doing business with Lehman experienced a higher probability of failure, and the liquidity of the stocks that they were trading declined. Coval and Stafford (2007) find that stocks sold by mutual funds that experience extreme outflows drop significantly in price during the outflow period and earn abnormally high expected returns over the next eighteen months. 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. Acharya, Lochstoer, and Ramadorai (2013) find that risk premia in commodity-futures markets are larger when broker-dealer balance sheets are shrinking.

can be “long-lived” and maximize constant absolute risk-aversion (CARA) utility over an infinite consumption stream, or they can be “short-lived” and maximize a mean-variance objective over changes in wealth in the next instant. Under both specifications, hedgers’ demand for insurance is independent of their wealth. On the other hand, because arbitrageurs maximize constant relative risk aversion (CRRA) utility over consumption, the supply of insurance depends on their wealth.

Arbitrageur wealth impacts equilibrium prices and allocations, and is the key state variable in our model. Solving for equilibrium amounts to solving a system of ordinary differential equations (ODEs) in wealth, with boundary conditions at zero and infinity. These ODEs include non-linear terms. Yet, their structure makes it possible to prove general analytical results across the entire parameter space, e.g., for all risk-aversion parameters of hedgers and arbitrageurs. In the case where hedgers are short-lived, we show that a solution exists and we characterize how it depends on wealth and on model parameters. Moreover, in both the short-lived and long-lived cases, we characterize the behavior of the solution close to the boundaries.

Our analysis yields new insights on dynamic risk-sharing and asset pricing. We show that the risk aversion of arbitrageurs is the sum of their static CRRA coefficient and of a forward-looking component that reflects intertemporal hedging. The latter component makes the risk aversion of arbitrageurs dependent on parameters of the economy that affect equilibrium prices. For example, when hedgers are more risk-averse, arbitrageurs become endogenously more risk averse if their CRRA is smaller than one. This effect can be sufficiently strong to imply that more risk-averse hedgers may receive *less* insurance from arbitrageurs in equilibrium. Intuitively, when hedgers are more risk-averse, expected returns rise steeply following declines in arbitrageur wealth. This makes arbitrageurs with CRRA smaller than one willing to invest more conservatively, so to preserve wealth in bad states and earn the high returns.

On the asset-pricing side, we show that arbitrageurs generate endogenous risk, in the sense that changes in their wealth affect return variances and covariances through amplification and contagion mechanisms. Endogenous risk is small at both extremes of the wealth distribution: when wealth is close to zero this is because arbitrageurs hold small positions and hence have a small impact on prices, and when wealth is close to infinity this is because prices are insensitive to changes in wealth. The dependence of endogenous risk on arbitrageur wealth can give rise to hump-shaped patterns of variances, covariances and correlations. It can also cause risk premia, defined as expected returns in excess of the riskless asset, to increase with arbitrageur wealth for small values of wealth, even though Sharpe ratios decrease. We show that risk premia always exhibit this pattern when arbitrageurs’ CRRA coefficient is small, and can exhibit it for larger values as well provided that hedgers are sufficiently risk-averse. In a similar spirit, we show that risk premia can be larger if the arbitrageurs' CRRA coefficient is smaller—precisely because endogenous risk is larger.

Additional asset-pricing results concern liquidity risk and its relationship with expected returns. A large empirical literature has documented that liquidity varies over time and in a correlated manner across assets within a class. Moreover, aggregate liquidity appears to be a priced risk factor and carry a positive premium: assets that underperform the most during times of low aggregate liquidity earn higher expected returns than assets with otherwise identical characteristics.[2] We map our model to that literature by defining liquidity based on the impact that hedgers have on prices. We show that liquidity is lower for assets with more volatile cashflows. It also decreases following losses by arbitrageurs, and this variation is common across assets.

Expected returns in our cross-section of assets are proportional to the covariance with the portfolio of arbitrageurs, which is the single priced risk factor. That factor may be hard to measure empirically as the portfolio of arbitrageurs is unobservable. We show, however, that aggregate liquidity captures that factor. Indeed, because arbitrageurs sell a fraction of their portfolio following losses, assets that covary the most with their portfolio suffer the most when liquidity decreases. Thus, an asset's covariance with aggregate liquidity is proportional to its covariance with the portfolio of arbitrageurs. On the other hand, the covariances between an asset's liquidity and aggregate liquidity or return would not explain expected returns as well. This is because they are proportional to the volatility of an asset's cashflows rather than to the asset's covariance with the arbitrageurs' portfolio. The covariance between an asset's return and other proxies of arbitrageur wealth, such as the leverage of financial intermediaries, used in recent empirical papers, would also capture the true priced risk factor.[3]

We finally characterize when the long-run stationary distribution of arbitrageur wealth is non-degenerate, and show that it can be bimodal. The stationary distribution can be non-degenerate because arbitrage activity is self-correcting: when wealth drops, the arbitrageurs' future expected returns increase causing wealth to grow faster, and vice-versa. The stationary density becomes bimodal when hedgers are sufficiently risk-averse. Indeed, because insurance provision in that case

[2]Chordia, Roll, and Subrahmanyam (2000), Hasbrouck and Seppi (2001), and Huberman and Halka (2001) document the time-variation of liquidity in the stock market and its correlation across stocks. 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 risk factor in the stock market and carries a positive premium. Sadka (2010) and Franzoni, Nowak, and Phalippou (2012) find similar results 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.

[3]Adrian, Etula, and Muir (2014) and He, Kelly, and Manela (2017) find that a single risk factor based on intermediary leverage can price a large cross-section of assets. As we argue in Section 5, our model is exactly consistent with this finding, hence suggesting that an explanation with minimal frictions may be possible.

is more profitable, arbitrageur wealth grows fast and large values of wealth can be more likely in steady state than 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 are more likely than intermediate values.

We see our work as bridging three relatively distinct streams of theoretical literature: on liquidity risk, on intermediary asset pricing, and on consumption-based asset pricing with heterogeneous agents. The first stream focuses on the pricing of liquidity risk in the cross-section of assets. In Holmstrom and Tirole (2001), firms avoid assets whose return is low when financial constraints are severe, and these assets offer high expected returns in equilibrium. The covariance between asset returns and liquidity (less severe constraints) is exogenous. It is instead endogenous in our model because prices depend on arbitrageur wealth, and this endogeneity is key for our results on a priced liquidity factor. 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: low liquidity and low prices are endogenous symptoms of low arbitrageur wealth.

The second stream links intermediary capital to liquidity and asset prices. 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.[4] 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 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 (2014), arbitrageurs are more efficient holders of productive capital and can trade a risky claim to that capital with other investors. The long-run stationary distribution of arbitrageur wealth can have a bimodal density. A key difference

[4] In Gromb and Vayanos (2017), arbitrage spreads are positively related to the spreads' sensitivity to arbitrageur wealth because both characteristics are positively related to cashflow volatility and convergence horizon.

with all of the above papers is that we derive the effects of arbitrage capital without imposing any constraints or contracting frictions.

Perhaps the closest papers to ours within the second stream of literature are Xiong (2001) and Kyle and Xiong (2001). In both papers, arbitrageurs with logarithmic utility over consumption can trade with exogenous long-term traders and noise traders over an infinite horizon.[5] The liquidity provided by arbitrageurs is increasing in their wealth, and asset volatilities are hump-shaped. Relative to these papers, we derive the demand of all traders from optimizing behavior and consider a general number of risky assets. We also show results analytically (rather than via numerical examples), and do so for general parameter values within which logarithmic preferences are a restrictive special case.

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), Longstaff and Wang (2012), Basak and Pavlova (2013), Chabakauri (2013), Garleanu and Panageas (2015), and Ehling and Heyerdahl-Larsen (2017). In these papers, agents have CRRA utility and differ in their risk aversion. As the wealth of the less risk-averse agents increases, Sharpe ratios decrease, and this can cause volatilities and correlations to be hump-shaped.[6] In contrast to these papers, we assume that only one set of agents has wealth-dependent risk aversion, and we fix the riskless rate.

We proceed as follows. In Section 2 we present the model. In Sections 3 and 4 we solve for equilibrium assuming that the risky assets are in zero supply. In Section 5 we explore the implications of our model for liquidity risk. In Section 6 we show that our main results extend to positive supply. Section 7 concludes, and all proofs are in the Appendix.

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 asset whose instantaneous return is constant over time and equal to $r > 0$. This return is exogenous in our model and could be derived from a linear and

[5] Isaenko (2008) studies a related model in which long-term traders maximize CARA utility and there are transaction costs.

[6] Longstaff and Wang (2012) show that the hump-shaped pattern extends to expected excess returns. Garleanu and Pedersen (2011) also find hump-shaped volatilities and expected returns. These findings, however, are shown via numerical examples rather than general proofs. See also Liu, Lu, Sun, and Yan (2017) for a model in which arbitrageurs render anomalies that they discover endogenously more correlated and this endogenous risk is hump-shaped in wealth.

riskless production technology. There are N risky assets with cashflows

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

where $\bar{D}$ is a constant $N \times 1$ vector, σ is a constant and invertible $N \times N$ matrix, and $\top$ denotes transpose. The cashflows (2.1) are *i.i.d.* The *i.i.d.* assumption is for simplicity; we can introduce persistence without significant changes to our analysis. We denote by S_t the $N \times 1$ vector of risky-asset prices at time t, and by s the $N \times 1$ vector consisting of asset supplies measured in terms of number of shares. The prices S_t are determined endogenously in equilibrium. We set $\Sigma \equiv \sigma^\top \sigma$.

There are two sets of agents, hedgers and arbitrageurs. Each set forms a continuum with measure one. Hedgers receive a random endowment $u^\top dD_t$ at $t + dt$, where u is a constant $N \times 1$ vector. Because the endowment is correlated with the risky assets' cashflows, it can be hedged by trading in these assets. We consider two specifications for hedgers' preferences. Under both specifications, the hedgers' wealth does not affect their risk aversion and demand for insurance. We intentionally simplify the model in this respect, so that we can focus on the supply of insurance, which is time-varying because of the wealth-dependent risk aversion of arbitrageurs.

- **Specification 1: Long-lived hedgers.** Hedgers maximize negative-exponential utility over intertemporal consumption:

$$-\mathrm{E}_t \left(\int_t^\infty e^{-\frac{\alpha}{r}\bar{c}_{t'}} e^{-\bar{\rho}(t'-t)} dt' \right), \tag{2.2}$$

  where $\bar{c}_{t'}$ is consumption at $t' \geq t$, $\frac{\alpha}{r}$ is the coefficient of absolute risk aversion, and $\bar{\rho}$ is the subjective discount rate.

- **Specification 2: Short-lived hedgers.** Hedgers maximize a mean-variance objective over instantaneous changes in wealth:

$$\mathrm{E}_t(dv_t) - \frac{\alpha}{2}\mathrm{Var}_t(dv_t), \tag{2.3}$$

  where dv_t is the change in wealth between t and $t + dt$, and α is a risk-aversion coefficient.

The risk-aversion coefficient $\frac{\alpha}{r}$ under Specification 1 is over consumption, and it yields a risk-aversion coefficient α over wealth, same to that under Specification 2. The interpretation of hedgers under Specification 1 is straightforward: they are infinitely lived agents. Under Specification 2, hedgers can instead be interpreted as overlapping 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. If preferences over consumption are described by the VNM utility U, this yields the objective (2.3) with the risk-aversion coefficient $\alpha = -\frac{U''(\bar{v})}{U'(\bar{v})}$.

Under the overlapping generations interpretation, Specification 2 introduces the friction that future generations of hedgers cannot trade with the current generation. Markets are hence incomplete, although they are complete for the current generation of hedgers and for arbitrageurs because the number N of risky assets is equal to the number of Brownian motions.[7] Under Specification 1, markets are complete for all hedgers and arbitrageurs. While generating a form of incompleteness, Specification 2 has the advantage of being more tractable. We refer to Specification 1 as long-lived hedgers, and to Specification 2 as short-lived hedgers.

Arbitrageurs maximize power utility over intertemporal consumption. When the coefficient γ of relative risk aversion is different than one, the arbitrageurs' objective at time t is

$$\mathrm{E}_t\left(\int_t^\infty \frac{c_{t'}^{1-\gamma}}{1-\gamma} e^{-\rho(t'-t)} dt'\right), \tag{2.4}$$

where $c_{t'}$ is consumption at $t' \geq t$ and ρ is the subjective discount rate. When $\gamma = 1$, the objective becomes

$$\mathrm{E}_t\left(\int_t^\infty \log(c_{t'}) e^{-\rho(t'-t)} dt'\right). \tag{2.5}$$

We assume that $\rho > r$. As we explain in Section 3.3, this assumption ensures that arbitrageurs do not accumulate infinite wealth over time, in which case wealth effects become irrelevant.

In Sections 3 and 4 we solve for equilibrium assuming that the risky assets are in zero supply ($s = 0$). With zero supply, the payoffs that hedgers derive at $t + dt$ from their risky positions at t are the opposite of those derived by arbitrageurs. The same payoffs can be derived through suitable positions in "short-maturity" assets that pay the same cashflows as the "long-maturity" assets at $t + dt$ and zero thereafter, and are in zero supply. This is because the diffusion matrix σ of cashflows has full rank N. Therefore, an equilibrium with long-maturity assets in zero supply yields the same risk-sharing, market prices of risk, and wealth dynamics, as one with short-maturity

[7] Market incompleteness prevents future generations of hedgers from hedging against the risk that when they are born arbitrageur wealth is low and the cost of insurance is high.

assets in zero supply. Because price dynamics with short-maturity assets are simpler than with long-maturity assets, we derive the latter equilibrium first, in Section 3. In Section 4 we confirm that the two equilibria are equivalent, and derive the long-maturity assets' prices, expected returns, volatilities, and correlations. Although supply is zero, there is aggregate risk because of the hedgers' endowment, and risk premia are non-zero and time-varying.

We allow supply to be positive in Section 6. We show that when hedgers are long-lived, risk-sharing and asset prices are the same as in the zero-supply equilibrium derived in Sections 3 and 4, provided that we replace u by $s+u$. That is, only the aggregate of the supply u coming from hedgers and the supply s coming from issuers matters, and not the relative composition. When hedgers are short-lived, this equivalence does not hold. Yet, key aspects of asset-price behavior derived in Sections 3 and 4 generalize.

For zero supply our model could represent futures markets, with the assets being futures contracts and the arbitrageurs being the speculators. It could also represent the market for insurance against aggregate risks, e.g., weather or earthquakes, with the assets being insurance contracts and the arbitrageurs being the insurers. For positive supply, our model could represent stock or bond markets, with the arbitrageurs being hedge funds or other agents absorbing demand or supply imbalances.

3 Equilibrium with Short-Maturity Assets

A new set of N short-maturity assets can be traded at each time t. The assets available at t pay dD_t at $t+dt$ and zero thereafter. We denote by $\pi_t dt$ the $N\times 1$ vector of prices at which the assets trade at t, and by $dR_t \equiv dD_t - \pi_t dt$ the $N\times 1$ vector of returns that the assets earn between t and $t+dt$. Equation (2.1) implies that the instantaneous expected returns of the short-maturity assets are

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

and the instantaneous covariance matrix of returns 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. Note also that dR_t is a return per share rather than per dollar invested: computing the return per dollar would require dividing dR_t by the price $\pi_t dt$. When using dollar rather than share returns in the rest of this paper, we will be mentioning that explicitly.

3.1 Optimization

Consider first the optimization problem of a long-lived hedger. The hedger's budget constraint is

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

where x_t is the hedger's position in the risky assets at time t and v_t is the hedger's wealth. 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, the third term is the endowment (with u characterizing the endowment's sensitivity to asset returns), and the fourth term is consumption. We solve the hedger's optimization problem using dynamic programming and conjecture the value function

$$V(v_t, w_t) = -e^{-[\alpha v_t + F(w_t)]}, \tag{3.4}$$

where $F(w_t)$ is a scalar function of w_t. The hedger's value function over wealth has the same negative-exponential form as the utility function over consumption, with the risk-aversion coefficient being α rather than $\frac{\alpha}{r}$. In addition, the value function depends on the wealth of arbitrageurs since the latter affects asset prices π_t. Arbitrageur wealth is the only state variable in our model.

Proposition 3.1 *Given the value function (3.4), the optimal policy of a long-lived hedger at time t is to consume*

$$\bar{c}_t = rv_t + \frac{r}{\alpha}\left[F(w_t) - \log(r)\right] \tag{3.5}$$

and hold a position

$$x_t = \frac{\Sigma^{-1}(\bar{D} - \pi_t)}{\alpha} - u - \frac{F'(w_t) y_t}{\alpha} \tag{3.6}$$

in the risky assets.

The hedger's optimal demand for the risky assets consists of three components, which correspond to the three terms in the right-hand side of (3.6). The first term is a standard mean-variance demand. It 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 covariance matrix Σ of asset returns times the vector $\bar{D} - \pi_t$ of expected returns. The second term is a demand to hedge endowment risk. It 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 $u_n > 0$ yields a high payoff when dD_{nt} is low, which when is the endowment is also low. The third term is an intertemporal hedging demand as in Merton (1971). Changes in arbitrageur wealth, the only state variable in our model, affect the terms at which hedgers can obtain insurance, and must be hedged against. Intertemporal hedging is accomplished by holding a portfolio with weights proportional to the sensitivity of arbitrageur wealth to asset returns. That sensitivity is simply the portfolio y_t of arbitrageurs. Hence, the intertemporal hedging demand is a scaled version of y_t, as the third term in (3.6) confirms.

When the hedger is short-lived, the budget constraint (3.3) does not include consumption, and the hedger's optimal demand for the risky assets does not include the intertemporal hedging component. The other two terms in (3.6), however, remain the same. Hence, the case of a short-lived hedger can be nested into that of a long-lived hedger by setting the function $F(w_t)$ to zero.

Proposition 3.2 *The optimal policy of a short-lived hedger at time t is to hold a position*

$$x_t = \frac{\Sigma^{-1}(\bar{D} - \pi_t)}{\alpha} - u \tag{3.7}$$

in the risky assets.

Consider next the optimization problem of an arbitrageur. The arbitrageur's budget constraint is

$$dw_t = rw_t dt + y_t^\top (dD_t - \pi_t dt) - c_t dt, \tag{3.8}$$

where y_t is the arbitrageur's position in the risky assets at time t and w_t is the arbitrageur's wealth. The first term in the right-hand side of (3.8) 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 aggregate wealth of all arbitrageurs since the latter affects asset prices π_t. In equilibrium own wealth and aggregate wealth coincide because all arbitrageurs are symmetric and in measure one. For the purposes of optimization, however, we need to make the distinction. We reserve the notation w_t for aggregate wealth and denote own wealth by $\hat{w}_t$. We likewise use (c_t, y_t) for aggregate consumption and position in the assets, and denote own consumption and position by $(\hat{c}_t, \hat{y}_t)$. We conjecture the value function

$$V(\hat{w}_t, w_t) = q(w_t)\frac{\hat{w}_t^{1-\gamma}}{1-\gamma} \tag{3.9}$$

for $\gamma \neq 1$, and

$$V(\hat{w}_t, w_t) = \frac{1}{\rho}\log(\hat{w}_t) + q_1(w_t) \tag{3.10}$$

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 $\gamma = 1$.

Proposition 3.3 *Given the value function (3.9) and (3.10), the optimal policy of an arbitrageur at time t is to consume*

$$\hat{c}_t = q(w_t)^{-\frac{1}{\gamma}}\hat{w}_t \tag{3.11}$$

and hold a position

$$\hat{y}_t = \frac{\Sigma^{-1}(\bar{D} - \pi_t)}{\frac{\gamma}{\hat{w}_t}} + \frac{q'(w_t)y_t}{\frac{\gamma}{\hat{w}_t}q(w_t)} \tag{3.12}$$

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 aggregate arbitrageur wealth w_t. The arbitrageur's optimal demand for the risky assets consists of two components, which correspond to the two terms in the right-hand side of (3.12) and are analogous to those for hedgers. The first term 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 term is an intertemporal hedging demand. The arbitrageur hedges against changes in aggregate arbitrageur wealth since these affect asset prices. Intertemporal hedging is accomplished by holding a portfolio with weights proportional to the sensitivity of aggregate arbitrageur wealth to asset returns. Since this sensitivity is the aggregate portfolio y_t of arbitrageurs, the intertemporal hedging demand is a scaled version of y_t.

3.2 Equilibrium Characterization and Existence

Since in equilibrium all arbitrageurs are symmetric and in measure one, their aggregate position coincides with each arbitrageur's position and the same is true for wealth. Setting $\hat{y}_t = y_t$ and $\hat{w}_t = w_t$ in (3.12), we find that the aggregate position of arbitrageurs is

$$y_t = \frac{\Sigma^{-1}(\bar{D} - \pi_t)}{A(w_t)}, \tag{3.13}$$

where

$$A(w_t) \equiv \frac{\gamma}{w_t} - \frac{q'(w_t)}{q(w_t)}. \tag{3.14}$$

Arbitrageurs' investment in the tangent portfolio is scaled by $A(w_t)$, which can be interpreted as a coefficient of *dynamic risk aversion*. It 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 and hence reflects dynamic considerations. Suppose, for example, that $q(w_t)$ is decreasing, a property which holds for $\gamma < 1$, as we show in Theorem 3.1. Equation (3.12) then implies that the intertemporal hedging demand is negative and lowers the arbitrageurs' position. The negative hedging demand is reflected in (3.13) through a larger coefficient of dynamic risk aversion $A(w_t)$. In Section 3.3.2 we provide economic intuition for the sign of $q'(w_t)$ and of the intertemporal hedging demand.

A similar calculation can be made for hedgers using the market-clearing equation

$$x_t + y_t = 0. \tag{3.15}$$

Setting $y_t = -x_t$ in (3.6), we find that the aggregate position of hedgers is

$$x_t = \frac{\Sigma^{-1}(\bar{D} - \pi_t) - \alpha u}{\alpha - F'(w_t)}. \tag{3.16}$$

Substituting (3.13) and (3.16) into (3.15), we find that asset prices π_t are

$$\pi_t = \bar{D} - \frac{\alpha A(w_t)}{\alpha + A(w_t) - F'(w_t)} \Sigma u. \tag{3.17}$$

Substituting (3.17) back into (3.13), we find that the arbitrageurs' position in the risky assets in equilibrium is

$$y_t = \frac{\alpha}{\alpha + A(w_t) - F'(w_t)} u. \tag{3.18}$$

Arbitrageurs buy a fraction of the portfolio u, which is the portfolio that hedgers want to sell to hedge their endowment risk. They buy a larger fraction of u, hence supplying more insurance to hedgers, when their coefficient of dynamic risk aversion $A(w_t)$ is smaller and the hedgers' risk-aversion coefficient α is larger. The degree of insurance supplied by arbitrageurs also depends on the hedgers' intertemporal hedging demand, as reflected through the function $F'(w_t)$. When hedgers are short-lived, $F(w_t) = 0$ and hence $F'(w_t) = 0$. When instead they are long-lived, our numerical solutions indicate that $F'(w_t) > 0$. Therefore, the hedgers' intertemporal hedging motive makes them demand more insurance from arbitrageurs. The intuition is that hedgers seek to hedge against the event that arbitrageurs will become poorer in the future because in that event insurance will be supplied to them at worse terms. Arbitrageurs become poorer when assets they have bought from hedgers under-perform. An individual hedger can hedge against that event by selling an even larger amount of those assets to arbitrageurs. As a consequence, the intertemporal hedging demand makes the hedgers collectively demand more insurance from arbitrageurs.

Equation (3.17) implies that expected asset returns $\bar{D} - \pi_t$ are proportional to the covariance with the portfolio u, which is the single priced risk factor in our model. The risk premium $\frac{\alpha A(w_t)}{\alpha + A(w_t) - F'(w_t)}$ of that factor depends on arbitrageur wealth, 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 depends on their wealth. Using (3.17) and (3.18), we find that the Sharpe ratio is

$$SR_t \equiv \frac{y_t^\top (\bar{D} - \pi_t)}{\sqrt{y_t^\top \Sigma y_t}} = \frac{\alpha A(w_t)}{\alpha + A(w_t) - F'(w_t)} \sqrt{u^\top \Sigma u}. \tag{3.19}$$

Substituting the hedgers' optimal policy from Proposition 3.1 into their Bellman equation (Equation (A.6), derived in the proof of Proposition 3.2), and using the dynamics that arbitrageur wealth follows in equilibrium, we can derive an ordinary differential equation (ODE) for the function $F(w_t)$. Following the same procedure for arbitrageurs, we can derive an ODE for the function $q(w_t)$. To state these ODEs and subsequent results, we define the parameter $z > 0$ by

$$z \equiv \frac{\alpha^2 u^\top \Sigma u}{2(\rho - r)}. \tag{3.20}$$

This parameter is larger when the hedgers' risk-aversion coefficient α or endowment variance $u^\top \Sigma u$ are larger, or the arbitrageurs' subjective discount rate ρ is smaller, or the riskless rate r is larger.

Proposition 3.4 *In equilibrium, the function $q(w_t)$ solves the ODE*

$$1 = \frac{q(w_t)^{-\frac{1}{\gamma}} - r}{\rho - r} A(w_t) w_t - \frac{z\left(A'(w_t) + A(w_t)^2\right)}{[\alpha + A(w_t) - F'(w_t)]^2}. \tag{3.21}$$

The function $F(w_t)$ is equal to zero when hedgers are short-lived, and solves the ODE

$$\begin{aligned} 1 = & \frac{rF(w_t) - r\log(r) - \alpha u^\top \bar{D} + \rho - \bar{\rho}}{\rho - r} + \frac{q(w_t)^{-\frac{1}{\gamma}} - r}{\rho - r} F'(w_t) w_t \\ & - \frac{z\left\{F''(w_t) - A(w_t)\left[2\alpha + A(w_t) - 2F'(w_t)\right]\right\}}{\left[\alpha + A(w_t) - F'(w_t)\right]^2} \end{aligned} \tag{3.22}$$

when they are long-lived.

Solving for equilibrium when hedgers are short-lived amounts to solving the ODE (3.21) with the function $F(w_t)$ set to zero. That ODE involves the functions $q(w_t)$ and $A(w_t)$. Since, however, $A(w_t)$ depends on $q(w_t)$ and $q'(w_t)$, as described in (3.14), (3.21) can be written as a second-order ODE in the single function $q(w_t)$. That ODE involves non-linear terms in both $q(w_t)$ and $q'(w_t)$. Solving for equilibrium when hedgers are long-lived amounts to solving the ODEs (3.21) and (3.22). These can be written as a system of second-order non-linear ODEs in $q(w_t)$ and $F(w_t)$. We derive boundary conditions for the two ODEs through a small set of economic properties which we assume should hold when w_t goes to zero and to infinity. We next state and motivate these properties.

When w_t goes to zero and to infinity, expected asset returns should converge to finite limits: if the limits were infinite, insurance would become infinitely costly, and hedgers would not only refrain from buying it but would also be willing to supply it. Since expected returns converge to finite limits and hedgers have negative exponential utility, their value function should also converge to finite limits, holding their wealth v_t constant. Equation (3.4) then implies that $F(w_t)$ should converge to finite limits when w_t goes to zero and to infinity.

That expected returns converge to finite limits when w_t goes to zero and to infinity does not imply that an arbitrageur's value function should do the same, holding the arbitrageur's own wealth $\hat{w}_t$ constant. Indeed, because arbitrageurs have power utility, their value function can become infinite when expected returns are large enough.[8] For w_t going to infinity, however, the finite limit of dollar expected returns is the riskless rate r, as arbitrageurs eliminate all risk premia. Moreover,

[8] Merton (1971) shows than when dollar return distributions are time-invariant and expected returns are finite, the value function remains finite under negative exponential utility but can become infinite under power utility.

because the arbitrageurs' subjective discount rate ρ exceeds r, their value function does converge to a finite limit, and so does $q(w_t)$.[9] The limit of $q(w_t)$ must further be positive: since $q(w_t)$ is the marginal utility of wealth of an arbitrageur with wealth $\hat{w}_t = 1$, and the arbitrageur can always invest in the riskless asset, $q(w_t)$ must exceed a positive bound. This is the boundary condition for $q(w_t)$ at infinity. The boundary condition at zero is through $A(w_t)$. Since arbitrageurs have power utility and risk premia are finite and non-zero, the arbitrageurs' position in the risky assets as a fraction of their wealth should converge to a finite non-zero limit (while converging to zero in absolute terms). Equation (3.13) then implies that $A(w_t)w_t$ should converge to a positive limit. Note that all the boundary conditions that we impose concern the existence of finite limits rather than the limits' exact values.

Theorem 3.1 provides a comprehensive analysis of the equilibrium when hedgers are short-lived, and a partial analysis when they are long-lived. In the short-lived case, we show that a solution to the ODE (3.21) with $F(w_t) = 0$ and the boundary conditions on $q(w_t)$ and $A(w_t)$ exists. We also characterize monotonicity properties of the solution and show that the limits at zero and infinity are uniquely determined. In the long-lived case, we have not been able to show existence and monotonicity properties. We show, however, that the limits at zero and infinity are the same as in the short-lived case.

Theorem 3.1 *When hedgers are short-lived, a solution to the ODE (3.21) with $F(w_t) = 0$ and positive limits of $A(w_t)w_t$ at zero and $q(w_t)$ at infinity exists. The solution has the following properties:*

- *The function $A(w_t)$ is decreasing.*
- $\lim_{w_t \to \infty} A(w_t)w_t = \gamma$ *and* $\lim_{w_t \to \infty} q(w_t) = \frac{1}{\left(r + \frac{\rho - r}{\gamma}\right)^{\gamma}}$.
- *If $\gamma < 1$, then $\frac{1}{w_t} > A(w_t) > \frac{\gamma}{w_t}$ and $q(w_t)$ is decreasing.*
- *If $\gamma > 1$, then $\frac{1}{w_t} < A(w_t) < \frac{\gamma}{w_t}$ and $q(w_t)$ is increasing.*
- *If $\gamma < K$, where $K < 1$ is the unique positive solution of*

$$G(\gamma) \equiv 1 - \gamma - \frac{\gamma}{z}\left(\frac{r\gamma}{\rho - r} + 1\right) = 0, \tag{3.23}$$

[9] The result that the value function under power utility is finite when dollar return distributions are time-invariant, expected returns are equal to the riskless rate r, and the subjective discount rate exceeds r is in Merton (1971).

then $\lim_{w_t \to 0} A(w_t)w_t = K$ *and* $\lim_{w_t \to 0} q(w_t) = \infty$.

- *If* $\gamma > K$, *then* $\lim_{w_t \to 0} A(w_t)w_t = \gamma$ *and* $\lim_{w_t \to 0} q(w_t) \in (0, \infty)$.

Suppose next that hedgers are long-lived, and that a solution to the system of ODEs (3.21) and (3.22) with positive limits of $A(w_t)w_t$ *at zero and* $q(w_t)$ *at infinity, and finite limits of* $F(w_t)$ *at zero and infinity, exists. Suppose additionally that* $F'(w_t)w_t$ *and* $F''(w_t)w_t^2$ *have (finite or infinite) limits at zero and infinity. The solution has the following properties:*

- *The limits of* $A(w_t)w_t$ *and* $q(w_t)$ *at zero and infinity are as in the case of short-lived hedgers.*
- $\lim_{w_t \to 0} F(w_t) = \log(r) + \frac{\alpha u^\top \bar{D} + \bar{\rho} - r - z(\rho - r)}{r}$ *and* $\lim_{w_t \to \infty} F(w_t) = \log(r) + \frac{\alpha u^\top \bar{D} + \bar{\rho} - r}{r}$.
- *The limits of* $[A(w_t) - F'(w_t)]w_t$ *at zero and infinity are the same as those of* $A(w_t)w_t$.

The basic idea of our existence proof is to start with a finite interval $[\epsilon, M]$ and show that there exists a unique solution to the ODE (3.21) with the limits of $A(w_t)w_t$ at zero and infinity imposed as boundary conditions at ϵ and M, respectively. We next show that when ϵ converges to zero and M to infinity, that solution converges to a solution over $(0, \infty)$. Our convergence proof uses the monotonicity of the solution with respect to ϵ and M, which in turn follows from a single-crossing property. Our construction yields a unique solution over $(0, \infty)$, although it does not rule out that other solutions (constructed differently) may exist. Uniqueness of our constructed solution allows us to examine how that solution moves in response to exogenous parameters: we perform the comparative statics on the solution over $[\epsilon, M]$ and take the limit when ϵ converges to zero and M to infinity. Our existence proof concerns only short-lived hedgers; when stating results on long-lived hedgers in the rest of this paper, we assume that a solution to the system of ODEs (3.21) and (3.22) as described in Theorem 3.1 exists.

In Section 3.3 we derive economic implications of the results shown in Theorem 3.1, as well as some additional properties. We examine how positions and returns depend on arbitrageur wealth (Section 3.3.1), how dynamic risk aversion differs from its static counterpart (Section 3.3.2), and what the long-run dynamics of arbitrageur wealth are (Section 3.3.3).

3.3 Equilibrium Properties

3.3.1 Wealth Effects

Theorem 3.1 shows that when hedgers are short-lived, an increase in the wealth w_t of arbitrageurs causes arbitrageur dynamic risk aversion $A(w_t)$ to decline. A decline in $A(w_t)$ results in more insurance supplied to hedgers: arbitrageur positions become 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. Expected asset returns, which reflect the cost of the insurance, become smaller in absolute value: less positive for positive elements of u and less negative for negative elements of u. 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.

Corollary 3.1 *When hedgers are short-lived, 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.

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

(iv) Lowers the arbitrageurs' Sharpe ratio.

The results of Corollary 3.1 appear consistent with the empirical findings of Kang, Rouwenhorst, and Tang (2016). That paper finds that commodity futures in which hedgers held long positions on average over the previous year earn negative expected returns. At the same time, commodity futures in which hedgers made a net purchase over the previous week (increasing their long position or reducing their short position) earn positive expected returns. In our model, hedgers hold long positions on average in assets n with $u_n < 0$, and these assets earn negative expected returns. At the same time, Corollary 3.1 implies that hedgers increase these long positions following an increase in arbitrageur wealth w_t, and the negative expected returns become less negative. Moreover, hedgers reduce their short positions in assets n with $u_n > 0$ following a decline in w_t, and the positive expected returns of these assets become more positive. Hence, net purchases of hedgers following shocks to w_t can on average (across both positive and negative shocks) be associated with positive expected returns. Hedging demand u and arbitrageur wealth w_t can thus generate relationships between positions and expected returns that are opposite and operate at different frequencies, consistent with the empirical evidence. In a similar spirit, Cheng and Xiong (2014) find that changes in the hedging demand of commodity hedgers do not appear to be the main driver of short-term changes in their positions.[10]

When hedgers are long-lived, changes in the wealth of arbitrageurs affect not only their dynamic risk aversion $A(w_t)$ but also the hedgers' intertemporal hedging demand, whose strength is captured by the function $F'(w_t) > 0$. Recall that long-lived hedgers seek to hedge against the event that arbitrageurs will become poorer in the future because in that event insurance will be supplied to them at worse terms. Moreover, because of that intertemporal hedging motive, hedgers demand more insurance. Our numerical solutions indicate that when arbitrageur wealth w_t increases, both $A(w_t)$ and $F'(w_t)$ decrease. In the case of $F'(w_t)$, this is because the terms of insurance become less sensitive to w_t for larger values of w_t.

The interplay between increased supply of insurance (lower $A(w_t)$) and declining demand for it (lower $F'(w_t)$) can give rise to non-monotonic patterns. Consider, for example, arbitrageur positions, which increase in absolute value when $A(w_t) - F'(w_t)$ declines, as shown in (3.18). When the hedgers' risk-aversion coefficient α is high, positions are hump-shaped in w_t: they increase in absolute value for small values of w_t as $A(w_t)$ declines, and decrease for larger values of w_t as $F'(w_t)$ declines. When instead α is low, the variation in $F'(w_t)$ is dominated by that in $A(w_t)$, and positions increase in absolute value for all values of w_t, as in Corollary 3.1.

Figure 1 illustrates the behavior of arbitrageur dynamic risk aversion and positions. The plots in the left show dynamic risk aversion $A(w_t)$. The plots in the middle show the position y_{nt} in an asset n, expressed as a fraction of the position u_n that hedgers want to hedge. All these plots concern the case where hedgers are long-lived. The plots in the right express positions relative to the case where hedgers are short-lived. The blue solid line represents the baseline case, which is the same in all plots. The plots at the top show that positions increase with w_t in the baseline case but become hump-shaped when the hedgers' risk-aversion coefficient α increases. For the larger value of α, arbitrageurs can over-insure hedgers, buying the full position that they want to hedge and holding an additional long position. Over-insurance introduces a large deviation relative to the case of short-lived hedgers, as the top-right figure shows. The plots at the bottom show that arbitrageurs provide less insurance to hedgers when their risk-aversion coefficient γ increases.

[10] Cheng and Xiong (2015) find that commodity hedgers scale down their positions when aggregate volatility (measured by VIX) increases. This is again suggestive of hedgers' positions changing in the short term for reasons unrelated to their hedging demand (which should increase when volatility increases): during volatile times, arbitrageurs may be becoming more constrained or risk-averse, an effect derived in Corollary 3.2 in the case $\gamma < 1$.

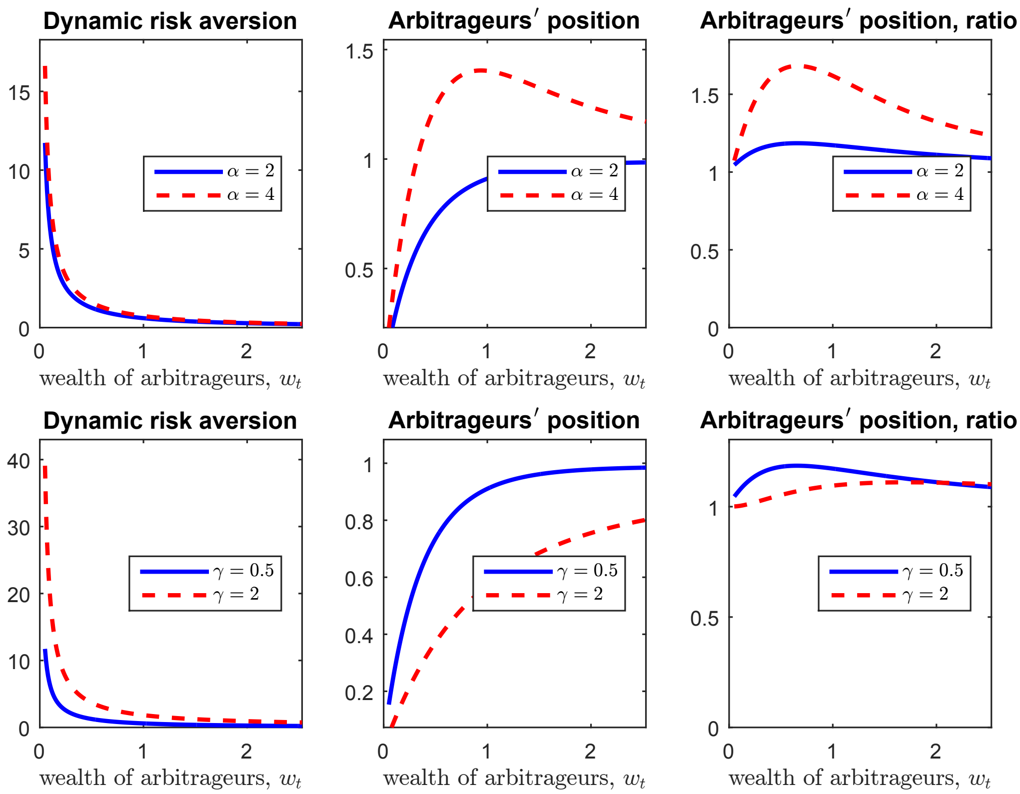

Figure 1: Arbitrageur dynamic risk aversion $A(w_t)$ and positions y_t as a function of arbitrageur wealth w_t. The plots in the left show $A(w_t)$. The plots in the middle show y_{nt} for an asset n, expressed as a fraction of the position u_n that hedgers want to hedge. All these plots concern the case where hedgers are long-lived. The plots in the right express positions relative to the case where hedgers are short-lived. All plots assume two symmetric and uncorrelated risky assets, and set the annualized standard deviation of the hedgers' endowment $\sqrt{u^\top \Sigma u}$ to 15%, the arbitrageurs' subjective discount rate ρ to 4%, and the riskless rate r to 2%. In the baseline case, represented by the blue solid line in all plots, the hedgers' risk-aversion coefficient α is set to 2 and the arbitrageurs' relative risk-aversion coefficient γ to 0.5. (By normalizing the hedgers' wealth to one, via a choice of numeraire, we can interpret α as their coefficient of relative risk aversion and $\sqrt{u^\top \Sigma u}$ as the annualized standard deviation of their endowment as a fraction of their wealth.) Under these choices, the parameters z and K are 2.25 and 0.59, respectively. The plots at the top examine the effect of raising α to 4, and the plots at the bottom examine the effect of raising γ to 2.

3.3.2 Dynamic Risk Aversion

Recall from (3.14) that the dynamic risk aversion $A(w_t)$ of arbitrageurs 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. In this section we examine, drawing on the results of Theorem 3.1, how dynamic and static risk aversion differ, or equivalently what the properties of intertemporal hedging demand are. We start with the case of short-lived hedgers, and examine how the results carry through to long-lived hedgers at the end of this section.

For $\gamma < 1$, the intertemporal hedging demand raises $A(w_t)$ above $\frac{\gamma}{w_t}$, while the opposite is true for $\gamma > 1$. The difference between the two cases lies in the behavior of the marginal utility of arbitrageur wealth, which is $q(w_t)$ for an arbitrageur with wealth $\hat{w}_t = 1$. For $\gamma < 1$, $q(w_t)$ is decreasing in w_t, meaning that the arbitrageur has higher marginal utility in states where aggregate arbitrageur wealth w_t is low. This is because in the low-w_t states expected returns are high, and hence the arbitrageur earns a high return on wealth. By seeking to preserve wealth in those states, so to earn the high return by investing it, the arbitrageur scales back his positions, behaving as more risk-averse than in the absence of intertemporal hedging. For $\gamma > 1$ instead, $q(w_t)$ is increasing in w_t, meaning that the arbitrageur has lower marginal utility in low-w_t states. This is because the high return on wealth in those states is associated with high utility of future consumption. Hence, it is discounted by low marginal utility, and the latter effect dominates for $\gamma > 1$.

When utility is logarithmic ($\gamma = 1$), the intertemporal hedging demand is zero, and hence $A(w_t) = \frac{1}{w_t}$. Although the intertemporal hedging demand raises $A(w_t)$ for $\gamma < 1$, $A(w_t)$ remains smaller than for $\gamma = 1$. Conversely, although the intertemporal hedging demand lowers $A(w_t)$ for $\gamma > 1$, $A(w_t)$ remains larger than for $\gamma = 1$. These comparisons do not imply, however, that the relative contribution of the intertemporal hedging demand to $A(w_t)$ is small; we next point out that it is large in some cases.

When w_t goes to infinity, $A(w_t) \approx \frac{\gamma}{w_t}$. Hence, $A(w_t)$ is driven purely by the static component, and the relative contribution of the intertemporal hedging demand converges to zero. The same result holds when w_t goes to zero if γ exceeds a threshold $K \in (0, 1)$. If instead $\gamma < K$, $A(w_t) \approx \frac{K}{w_t}$. Hence, the contribution of the intertemporal hedging demand to $A(w_t)$ is $\frac{K-\gamma}{w_t}$ in absolute terms and $\frac{K-\gamma}{K}$ in relative terms. If, in particular, γ is close to zero, $A(w_t)$ is almost purely driven by the intertemporal hedging demand. The difference between the cases $\gamma > K$ and $\gamma < K$ lies in the behavior of the marginal utility $q(w_t)$ when w_t goes to zero. If $\gamma > K$, then $q(w_t)$ converges to a finite limit and hence the ratio of $-\frac{q'(w_t)}{q(w_t)}$ to $\frac{\gamma}{w_t}$ must converge to zero. If instead $\gamma < K$, then $q(w_t)$ converges to infinity and hence $-\frac{q'(w_t)}{q(w_t)}$ can be larger than $\frac{\gamma}{w_t}$.[11]

[11]The region $\gamma < K$ in which $q(w_t)$ converges to infinity when w_t goes to zero is a subset of the region $\gamma < 1$ in which $q(w_t)$ is decreasing. The former region is also the one in which the value function of an arbitrageur facing the time-invariant expected returns that arise in equilibrium when $w_t = 0$ (no arbitrageurs) is infinite.

Equation (3.23) implies that the threshold K increases with the hedgers' risk-aversion coefficient α and endowment variance $u^\top \Sigma u$. Hence, when hedgers are more risk averse or their endowment is more volatile, arbitrageurs with $\gamma < K$ also become more risk averse for w_t close to zero. Theorem 3.2 generalizes these results to any $\gamma < 1$ and w_t, and shows that the opposite results hold for $\gamma > 1$.

Theorem 3.2 *Suppose that hedgers are short-lived. For any given level of arbitrageur wealth w_t, the following comparative statics hold:*

(i) An increase in the hedgers' risk-aversion coefficient α raises arbitrageur dynamic risk aversion $A(w_t)$ if $\gamma < 1$, and lowers it if $\gamma > 1$.

(ii) An increase in the hedgers' endowment variance $u^\top \Sigma u$ raises arbitrageur dynamic risk aversion $A(w_t)$ if $\gamma < 1$, and lowers it if $\gamma > 1$.

The intuition for Theorem 3.2 is that for larger values of α and $u^\top \Sigma u$, expected returns are more sensitive to changes in arbitrageur wealth w_t, rising more steeply when w_t declines. Hence, arbitrageurs with $\gamma < 1$ have a marginal utility that also rises more steeply following declines in w_t. This makes them even more willing to preserve wealth in the low-w_t states, and raises their dynamic risk aversion $A(w_t)$. Arbitrageurs with $\gamma > 1$, by contrast, have a marginal utility that drops more steeply following declines in w_t, and this lowers $A(w_t)$.

The comparative statics of $A(w_t)$ yield comparative statics for arbitrageur positions and Sharpe ratios. Most surprising among these is that arbitrageurs can supply *less* insurance to hedgers when the latter become more risk averse. Following an increase in α, supplying insurance becomes more profitable for arbitrageurs. If arbitrageur risk aversion included only the static component, which does not depend on α, arbitrageurs would therefore supply more insurance. Because of the intertemporal hedging demand, however, the dynamic risk aversion of arbitrageurs with $\gamma < 1$ increases. Moreover, this effect can dominate, inducing arbitrageurs to supply less insurance. Arbitrageurs with $\gamma \geq 1$ instead supply more insurance because their dynamic risk aversion decreases. When $\gamma \leq 1$, arbitrageurs' Sharpe ratio increases because both hedgers and arbitrageurs become more risk averse. The comparative statics with respect to $u^\top \Sigma u$ are along similar lines.

Corollary 3.2 *Suppose that hedgers are short-lived. For any given level of arbitrageur wealth w_t, the following comparative statics hold:*

(i) An increase in the hedgers' risk aversion coefficient α raises the arbitrageurs' position in absolute value when $\gamma \geq 1$ and lowers it when $\gamma < K$, $z < 1$ and w_t is small. It raises the arbitrageurs' Sharpe ratio when $\gamma \leq 1$.

(ii) An increase in the variance $u^\top \Sigma u$ of hedgers' endowment raises the arbitrageurs' position in absolute value when $\gamma > 1$ and lowers it when $\gamma < 1$. It raises the arbitrageurs' Sharpe ratio when $\gamma \leq 1$.

When hedgers are long-lived, an increase in α lowers the arbitrageurs' position in absolute value when $\gamma < K$, $z < 1$ and w_t is small.

The result that an increase in α can induce the arbitrageurs to supply less insurance to hedgers carries through to long-lived hedgers. This is because the asymptotic behavior of $A(w_t)$ for w_t close to zero is the same as with short-lived hedgers. Our numerical solutions indicate that the remaining comparative statics in Theorem 3.2 and Corollary 3.2 also extend to long-lived hedgers.

3.3.3 Stationary Distribution

We next derive the long-run dynamics of arbitrageur wealth.

Proposition 3.5 *If $z > 1$, then arbitrageur wealth has a long-run stationary distribution with density*

$$d(w_t) = \frac{\frac{[\alpha + A(w_t) - F'(w_t)]^2}{A(w_t)} \exp\left[\int_1^{w_t} \left(A(\hat{w}_t) - \frac{[\alpha + A(\hat{w}_t) - F'(\hat{w}_t)]^2}{zA(\hat{w}_t)}\right) d\hat{w}_t\right]}{\int_0^\infty \frac{[\alpha + A(w_t) - F'(w_t)]^2}{A(w_t)} \exp\left[\int_1^{w_t} \left(A(\hat{w}_t) - \frac{[\alpha + A(\hat{w}_t) - F'(\hat{w}_t)]^2}{zA(\hat{w}_t)}\right) d\hat{w}_t\right] dw_t} \tag{3.24}$$

over the support $(0, \infty)$. If $z < 1$, then wealth converges to zero in the long run.

Arbitrageur wealth has a non-degenerate stationary density if the parameter z defined in (3.20) exceeds one. That is, the hedgers' risk-aversion coefficient α and endowment variance $u^\top \Sigma u$ must be large enough relative to the difference between the arbitrageurs' subjective discount rate ρ and the riskless rate r. Note that this result is valid both when hedgers are short-lived and when they are long-lived.

The existence of a non-degenerate stationary density is related to the dynamics of arbitrageur wealth being *self-correcting*: when wealth drops, the arbitrageurs' future expected returns increase causing wealth to grow faster, and vice-versa. To explain the relationship and why the condition $z > 1$ is required, we recall the standard Merton (1971) portfolio optimization problem in which an infinitely-lived investor with CRRA coefficient γ can invest in a riskless rate r and in N risky assets whose returns have expectation given by a vector μ and covariance given by a matrix Σ. The investor's wealth converges to infinity in the long run when

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

i.e., 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 (Corollary 3.1). Using (3.19) to substitute for the arbitrageurs' Sharpe ratio, we can write (3.25) as

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

Transposing the result from the Merton problem thus suggests that there are three possibilities for the long-run dynamics of arbitrageur wealth. 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 small values, neither convergence occurs and wealth has a non-degenerate stationary density.

Since Theorem 3.1 shows that $A(w_t)$ converges to zero when w_t goes to infinity, (3.26) is satisfied for large values of w_t if $r > \rho$. Our assumption $\rho > r$ hence implies that (3.26) is violated for large w_t, and rules out that wealth converges to infinity in the long run. To examine whether (3.26) is satisfied for small values of w_t, we recall from Theorem 3.1 that $A(w_t)$ is of order $\frac{1}{w_t}$ for w_t close to zero, and that $\lim_{w_t \to 0} F'(w_t) w_t = 0$. Hence, (3.26) is satisfied for small w_t if $r + \frac{\alpha^2 u^\top \Sigma u}{2} > \rho$. This condition is equivalent to $z > 1$, which is exactly what Proposition 3.5 requires for a non-degenerate stationary density to exist. Proposition 3.6 computes the density in closed form when hedgers are short-lived and arbitrageurs have logarithmic utility ($\gamma = 1$).

Proposition 3.6 *Suppose that hedgers are short-lived, arbitrageurs have logarithmic utility ($\gamma = 1$), and $z > 1$. The stationary density $d(w_t)$:*

(i) Is decreasing in w_t if $z < \frac{27}{8}$.

(ii) Is bimodal in w_t otherwise: decreasing in w_t for $w_t \in (0, m_1)$, increasing in w_t for $w_t \in (m_1, m_2)$, and again decreasing in w_t for $w_t \in (m_2, \infty)$. The thresholds $m_1 < m_2$ are the two positive roots of

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

(iii) Shifts to the right in the monotone likelihood ratio sense when α or $u^\top \Sigma u$ increase.

The stationary density has two possible shapes. When z is not much larger than one, it is decreasing in arbitrageur wealth w_t, and so values of w_t close to zero are more likely than larger values. When instead z is sufficiently larger than one, the stationary density becomes bimodal, with the two maxima being zero and an interior point m_2 of the support. Values of w_t close to the maxima are more likely than intermediate values, meaning that the economy spends more time at these values than in the middle. The intuition is that when the hedgers' risk-aversion coefficient α and endowment variance $u^\top \Sigma u$ are large, arbitrageurs earn high expected returns for providing insurance, and their wealth grows fast. Therefore, large values of w_t can be 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.

Using the long-run stationary distribution, we can perform "unconditional" comparative statics. For example, rather than examining how the arbitrageurs' Sharpe ratio depends on α conditionally on w_t, we can examine how it depends on α unconditionally, in expectation over the stationary distribution of w_t. The unconditional comparative statics can differ from the conditional ones. For example, while an increase in α and $u^\top \Sigma u$ raises the conditional Sharpe ratio when $\gamma \leq 1$ (Corollary 3.2), it can lower its unconditional expectation. Intuitively, for larger values of α and $u^\top \Sigma u$, arbitrageur wealth grows faster, and its stationary density shifts to the right (Proposition 3.6). Therefore, while the conditional Sharpe ratio increases, its unconditional expectation can decrease because large values of wealth, which yield low Sharpe ratios, become more likely.

4 Equilibrium with Long-Maturity Assets

We conjecture that in equilibrium the price vector S_t of the long-maturity assets 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 denote by $dR_t \equiv dS_t + dD_t - rS_t dt$ the $N \times 1$ vector of returns that the long-maturity assets earn between t and $t + dt$ in excess of the riskless asset. Note as in the case of short-maturity assets, dR_t is a return per share rather than per dollar invested. Equations (2.1) and (4.1) imply that the instantaneous expected excess returns of the long-maturity assets are

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

and the instantaneous covariance matrix of returns is

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

4.1 Equivalence with Short-Maturity Assets

With long-maturity assets, the budget constraint (3.3) of a long-lived hedger becomes

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

where X_t is the hedger's position in the risky assets at time t. The budget constraint of a short-lived hedger is derived from (4.4) by excluding consumption. The budget constraint (3.8) of an arbitrageur becomes

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

where Y_t is the arbitrageur's position in the risky assets at time t. The market clearing equation (3.15) becomes

$$X_t + Y_t = 0. \tag{4.6}$$

Because the diffusion matrix σ of cashflows has full rank N, an equilibrium with long-maturity assets yields the same risk-sharing, market prices of risk, and wealth dynamics, as one with short-maturity assets.

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

(i) Choosing the price process S_t such that

$$\left(\sigma^\top\right)^{-1}(\bar{D} - \pi_t) = \left((\sigma_{St} + \sigma)^\top\right)^{-1}(\mu_{St} + \bar{D} - rS_t). \tag{4.7}$$

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

$$\sigma x_t = (\sigma_{St} + \sigma)X_t, \tag{4.8}$$

$$\sigma y_t = (\sigma_{St} + \sigma)Y_t. \tag{4.9}$$

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

Equations (4.8) and (4.9) construct positions of hedgers and arbitrageurs in the long-maturity assets so that the exposures to the underlying Brownian shocks are the same as with short-maturity assets. Equation (4.7) constructs a price process such that the market prices of the Brownian risks are also the same. Given this price process, risk exposures are optimal and clear the markets, because these properties also hold with short-maturity assets.

4.2 Asset Prices and Returns

The prices S_t of the long-maturity assets are a function of arbitrageur wealth w_t, which is the only state variable in our model. 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 (4.7), we can determine $S(w_t)$ up to an ODE for a scalar function.

Proposition 4.1 *The prices of the long-maturity assets are given by*

$$S(w_t) = \frac{\bar{D}}{r} + g(w_t)\Sigma u, \tag{4.10}$$

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

$$\frac{\alpha^2 u^\top \Sigma u}{2[\alpha + A(w_t) - F'(w_t)]^2} g''(w_t) + \left(r - q(w_t)^{-\frac{1}{\gamma}}\right) g'(w_t) w_t - r g(w_t) = \frac{\alpha A(w_t)}{\alpha + A(w_t) - F'(w_t)}. \tag{4.11}$$

The assets' expected excess returns are

$$\frac{\mathrm{E}_t(dR_t)}{dt} = \frac{\alpha A(w_t)}{\alpha + A(w_t) - F'(w_t)} \left[u^\top \Sigma u f(w_t) + 1\right] \Sigma u, \tag{4.12}$$

and the covariance matrix of returns is

$$\frac{\mathrm{Var}_t(dR_t)}{dt} = f(w_t) \left[u^\top \Sigma u f(w_t) + 2\right] \Sigma u u^\top \Sigma + \Sigma, \tag{4.13}$$

where

$$f(w_t) \equiv \frac{\alpha g'(w_t)}{\alpha + A(w_t) - F'(w_t)}. \tag{4.14}$$

The ODE (4.11) is linear in $g(w_t)$. The boundary conditions that we require are that $g(w_t)$ converges to finite limits at zero and infinity. As with the ODEs in Section 3, we only assume the existence of finite limits rather than the limits' exact values. Theorem 4.1 shows that a solution $g(w_t)$ to the ODE (4.11) exists when hedgers are short-lived, is negative and increasing in w_t, and converges to $-\frac{\alpha}{r}$ and to zero, respectively, when w_t goes to zero and to infinity. The theorem also shows that when hedgers are long-lived, the limits are as in the short-lived case.

Theorem 4.1 *When hedgers are short-lived, a solution to the ODE (4.11) with finite limits at zero and infinity exists. The solution has the following properties:*

- *The function $g(w_t)$ is negative and increasing.*
- $\lim_{w_t \to 0} g(w_t) = -\frac{\alpha}{r}$ *and* $\lim_{w_t \to \infty} g(w_t) = 0$.

When hedgers are long-lived, the limits of $g(w_t)$ at zero and infinity are the same as when they are short-lived, provided that a solution to the ODE (4.11) with finite limits exists.

Proposition 4.1 shows that asset prices (given in (4.10)) are the sum of two terms. The first term, $\frac{\bar{D}}{r}$, is the present value of the assets' expected cashflows $\bar{D}$, discounted at the riskless rate r. Prices would equal that present value if arbitrageurs had infinite wealth since they would then eliminate all risk premia, rendering the expected dollar returns on all assets equal to r. The second term, $g(w_t)\Sigma u$, reflects the risk premia arising from arbitrageur wealth w_t being finite. Consider an asset n that covaries positively with the portfolio u that hedgers want to sell, i.e., $(\Sigma u)_n > 0$. In the absence of arbitrageurs, that asset would trade at a discount relative to $\frac{\bar{D}_n}{r}$. Arbitrageurs cause that discount to decrease, and the more so the wealthier they are. Hence, the asset price $S_n = \frac{\bar{D}_n}{r} + g(w_t)(\Sigma u)_n$ increases in w_t and converges to $\frac{\bar{D}_n}{r}$ when w_t goes to infinity. Theorem 4.1 shows that $g(w_t)$ is indeed increasing and converges to zero at infinity.

Since changes in arbitrageur wealth w_t affect the prices of long-maturity assets, they also impact the assets' returns. Proposition 4.1 shows that the covariance matrix of asset returns (given in (4.13)) is the sum of a "fundamental" component Σ, driven purely by shocks to assets' underlying cashflows dD_t, and an "endogenous" component $f(w_t)\left[f(w_t)u^\top \Sigma u + 2\right] \Sigma u u^\top \Sigma$, introduced because cashflow shocks affect w_t which affects returns. Endogenous risk does not arise with short-maturity assets since their returns are risky only because of the payoff dD_t, which is not sensitive to changes in w_t.

The endogenous covariance between an asset pair (n, n') depends on whether the corresponding components of the vector Σu have the same or opposite signs. Suppose, for example, that $(\Sigma u)_n > 0$ and $(\Sigma u)_{n'} > 0$, in which case both assets would trade at a discount in the arbitrageurs' absence. An increase in arbitrageur wealth w_t causes the prices of both assets to increase, resulting in positive endogenous covariance. Suppose instead that $(\Sigma u)_n < 0$, in which case demand from hedgers would cause asset n to trade at a premium in the arbitrageurs' absence. Asset n's price would then drop following an increase in w_t, resulting in negative endogenous covariance with asset n'.

Endogenous risk is small at both extremes of the wealth distribution and larger in the middle. When arbitrageur wealth w_t is close to zero, arbitrageurs hold small positions in absolute terms (i.e., not as a fraction of w_t). Therefore, changes in w_t are small and have a small impact on prices. When instead w_t is close to infinity, arbitrageurs absorb the entire portfolio u that hedgers want to sell. Changes in w_t are hence larger, but prices are insensitive to those changes. These effects can be seen in the expression (4.14) for $f(w_t)$: this function is small for small w_t because $A(w_t)$ is large (resulting in small positions by arbitrageurs), and for large w_t because $g'(w_t)$ is small (low price sensitivity to w_t). Since the endogenous variance is larger in the middle than in the extremes, total variance can be hump-shaped in w_t. Total covariance and correlation can be hump-shaped or inverse hump-shaped depending on whether the endogenous covariance is positive or negative, respectively. Proposition 4.2 confirms the hump shapes (i.e., shows that there is only one hump) in the case where hedgers are short-lived and arbitrageurs have logarithmic utility.

Proposition 4.2 *The effects of a change in arbitrageur wealth w_t on the volatility of asset returns and on return covariance and correlation converge to zero when w_t goes to zero and to infinity. When hedgers are short-lived and arbitrageurs have logarithmic utility, an increase in w_t has:*

(i) A hump-shaped effect on the volatility of asset returns. The hump peaks at a value that is common to all assets.

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

(iii) The same hump-shaped effect on the correlation between the returns of assets n and n' if

$$\frac{(\Sigma u)_n(\Sigma u)_{n'}\Sigma_{nn} - (\Sigma u)_n^2\Sigma_{nn'}}{f(w_t)\left[f(w_t)u^\top\Sigma u + 2\right](\Sigma u)_n^2 + \Sigma_{nn}} + \frac{(\Sigma u)_n(\Sigma u)_{n'}\Sigma_{n'n'} - (\Sigma u)_{n'}^2\Sigma_{nn'}}{f(w_t)\left[f(w_t)u^\top\Sigma u + 2\right](\Sigma u)_{n'}^2 + \Sigma_{n'n'}} > 0, \quad (4.15)$$

and the opposite, i.e., inverse hump-shaped, effect if (4.15) holds in the opposite direction.

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 in covariance.

A hump-shaped pattern is possible for expected excess returns as well. This is more surprising because Corollary 3.1 shows that an increase in arbitrageur wealth w_t lowers the market prices of the Brownian risks, which are equal to expected excess returns per unit of risk exposure. Offsetting this effect, is that for w_t to the left of the volatility hump, an increase in w_t raises volatility because it raises endogenous risk. The latter effect can dominate and cause expected excess returns to increase with w_t for small values of w_t. Proposition 4.3 shows that when hedgers are short-lived, the latter effect always dominates for $\gamma < K$, and can dominate for larger values of γ as well provided that the parameter z is sufficiently large.

Proposition 4.3 *Suppose that hedgers are short-lived. For small arbitrageur wealth w_t, an increase in w_t raises the expected excess return of each asset in absolute value, if $\gamma < K$. If $\gamma > K$, then the same result holds provided that $z > \frac{\gamma\left(\frac{r}{\rho - r}\gamma + 1\right)}{\gamma + 1}$.*

Figure 2 plots the Sharpe ratios, expected excess returns, volatilities and correlations of long-maturity assets as a function of arbitrageur wealth for long-lived hedgers and for $\gamma = 0.5$ and 2. Consistent with Propositions 4.2 and 4.3, shown for short-lived hedgers, volatility and correlation are hump-shaped in arbitrageur wealth, and the hump-shape carries through to expected excess returns. The comparison across the two values of γ is also interesting. As one would expect, Sharpe ratios increase in γ. Expected excess returns, however, can be larger for the smaller value of γ. This is because when arbitrageurs are less risk-averse, they establish larger positions, and this generates more endogenous risk. As in Proposition 4.3, the effect of larger endogenous risk on expected excess returns can dominate that of smaller Sharpe ratios.

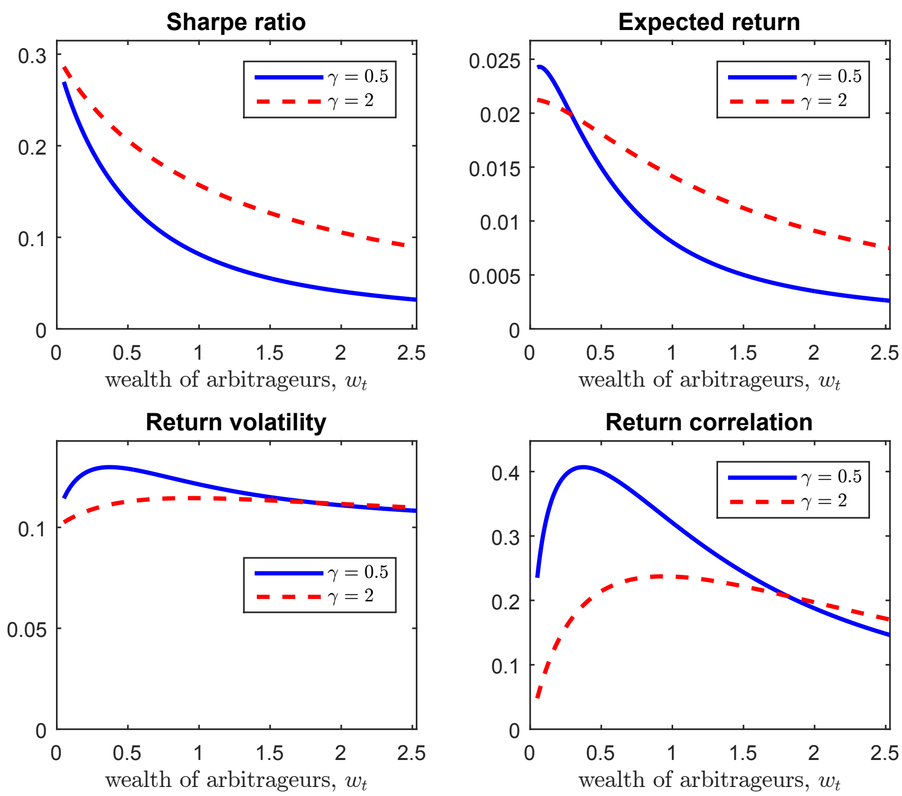

Figure 2: Assets' Sharpe ratios, expected excess returns, volatilities, and correlations as a function of arbitrageur wealth w_t. The plots assume two symmetric and uncorrelated risky assets, and set $\sqrt{u^\top \Sigma u} = 15\%$, $\alpha = 2$, $\rho = 4\%$ and $r = 2\%$. In the baseline case, represented by the blue solid line, γ is set to 0.5. The plots examine the effect of raising γ to 2. The baseline case is identical to that in Figure 1, which also examines the effect of raising γ to 2 (on dynamic risk aversion and positions).

5 Liquidity Risk

In this section we explore the implications of our model for liquidity risk. We assume long-maturity assets, as in the previous section, and 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

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

Defining illiquidity as price impact follows Kyle (1985). Kyle and Xiong (2001), Xiong (2001), and Johnson (2008) perform similar constructions to ours in asset-pricing settings by defining illiquidity as the derivative of price with respect to supply.

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

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

It converges to infinity when arbitrageur wealth w_t goes to zero, and to zero when w_t goes to infinity. When hedgers are short-lived, illiquidity is positive for all $w_t \in (0, \infty)$. When, in addition, arbitrageurs have logarithmic utility ($\gamma = 1$), illiquidity decreases in w_t.

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. Our numerical solutions indicate that illiquidity is a positive and decreasing function of wealth, a finding that Proposition 5.1 confirms in the case where hedgers are short-lived and arbitrageurs have logarithmic utility. The time-series variation of illiquidity is common across assets and corresponds to the term multiplying Σ_{nn} in (5.2). In the cross-section, illiquidity is higher for assets with more volatile cashflows. The dependence of illiquidity on an asset n is through Σ_{nn}, the asset's cashflow variance. 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.

We next 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

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

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 the illiquidity of an asset n 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.

When hedgers are short-lived and arbitrageurs have logarithmic utility ($\gamma = 1$), the proportionality coefficient $C^{\Lambda}(w_t)$ in the first relationship is negative, and those in the second and third relationships are positive and negative, respectively.

We next determine the link between liquidity-related covariances and expected returns. Equation (4.12) shows that the expected excess return of an 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, assets are priced by the portfolio of arbitrageurs, who are the marginal agents. Moreover, the covariance between asset returns and aggregate illiquidity identifies that portfolio perfectly. 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 to cashflow variance Σ_{nn}, which is only a component of $(\Sigma u)_n$. Thus, these covariances proxy for the true priced risk factor but imperfectly so.

Corollary 5.2 *In the cross-section of assets, expected excess returns are proportional to the covariance between returns and aggregate illiquidity. When hedgers are short-lived and arbitrageurs have logarithmic utility ($\gamma = 1$), the proportionality coefficient $\Pi^{\Lambda}(w_t)$ in this relationship is negative.*

The premium associated to the illiquidity risk factor is the expected excess return per unit of covariance with the factor. Hence, it coincides with the proportionality coefficient $\Pi^{\Lambda}(w_t)$ in the relationship between expected excess returns and covariance of returns with aggregate illiquidity.

Both the premium of the illiquidity risk factor and the covariance of returns with aggregate illiquidity vary over time in response to changes in arbitrageur wealth w_t. When w_t is low, illiquidity is high and highly sensitive to changes in wealth. Because of that effect, assets' covariance with illiquidity is large in absolute value when w_t is small, and decreases when w_t increases. Conversely, because the premium of the illiquidity risk factor is the expected excess return per unit of covariance, it is small in absolute value when w_t is small and increases when w_t increases.[12]

[12] Since the premium of the illiquidity risk factor depends on w_t, it can be viewed as a function of illiquidity itself, which is a monotone decreasing function of w_t.

Proposition 5.2 *The common component $C^{\Lambda}(w_t)$ of assets' covariance with aggregate illiquidity, and the premium $\Pi^{\Lambda}(w_t)$ of the illiquidity risk factor have the following properties:*

(i) $C^{\Lambda}(w_t)$ *converges to minus infinity when arbitrageur wealth w_t goes to zero, and to zero when w_t goes to infinity.*

(ii) $\Pi^{\Lambda}(w_t)$ *converges to zero when w_t goes to zero. When hedgers are short-lived,* $\Pi^{\Lambda}(w_t)$ *converges to minus infinity when w_t goes to infinity.*

Aggregate illiquidity explains expected returns perfectly in our model because it is a monotone function of arbitrageur wealth w_t. Hence, any other monotone function of w_t would also have this property. Recent empirical papers on intermediary asset pricing, such as Adrian, Etula, and Muir (2014) and He, Kelly, and Manela (2017), have used the leverage of specific groups of financial intermediaries as a risk factor, and have shown that it can price a large cross-section of assets. The findings of these papers are exactly consistent with our model, insofar as leverage is a monotone function of w_t.[13] Our modelling approach suggests that these findings may be explained even with minimal frictions, e.g., no borrowing constraints. On the other hand, introducing frictions may give better guidance on which function of w_t to use as a risk factor.

6 Positive Supply

Our analysis so far assumes that the long-maturity assets are in zero supply ($s = 0$). Proposition 6.1 shows that this assumption is without loss of generality when hedgers are long-lived: risk-sharing and asset prices are the same as when assets are in zero supply, provided that we replace u by $s+u$. The intuition is that the stream of random endowments $u^{\top} dD_t$ that a long-lived hedger receives over time is equivalent to an endowment of u shares in the long-maturity assets. Thus, hedgers generate a supply u, which is added to the supply s coming from the asset issuers. If instead asset issuers generate no supply and hedgers generate $s+u$, then hedgers reduce their demand by s. Since

[13] A natural measure of leverage in our model is the total risk exposure of arbitrageurs as a fraction of their wealth. Since the exposure of arbitrageurs to the Brownian risks is $\sigma y_t = (\sigma_{St} + \sigma)Y_t$, (3.13) implies that leverage is

$$\frac{1^{\top}\sigma y_t}{w_t} = \frac{\alpha 1^{\top}\sigma u}{\alpha w_t + A(w_t)w_t - F'(w_t)w_t}.$$

Lemma A.8, stated and proven in the Appendix, shows that this measure of leverage is decreasing in w_t when hedgers are short-lived and $\gamma \leq 1$. Our numerical solutions suggest that this result holds more generally. A countercyclical leverage is consistent with the empirical finding of He, Kelly, and Manela (2017), and the theory of He and Krishnamurthy (2013) and Brunnermeier and Sannikov (2014), but is inconsistent with Adrian, Etula, and Muir (2014).

the reduction in hedgers' demand exactly offsets the reduction in asset issuers' supply, equilibrium prices remain the same and so does risk-sharing.

Proposition 6.1 *Suppose that hedgers are long-lived and long-maturity assets are in positive supply s. If (S_t, X_t, Y_t) is an equilibrium, then $(S_t, X_t - s, Y_t)$ is an equilibrium when assets are in zero supply ($s = 0$) and u is replaced by $s + u$. In both equilibria, the exposures of hedgers and arbitrageurs to the Brownian shocks and the prices of the assets are the same.*

When hedgers are short-lived, the equivalence between positive-supply and zero-supply equilibria does not hold. This is because any given short-lived hedger receives a random endowment only in the next instant, so that endowment is not equivalent to one in the long-maturity assets. Prices and expected returns, however, have the same general form as with zero supply.

Proposition 6.2 *Suppose that hedgers are short-lived and that long-maturity assets are in positive supply s. Then, the prices of the assets are given by*

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

where the scalar function $g(w_t)$ solves the ODE (C.44). The assets' expected excess returns are

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

where

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

The price of an asset n, given in (6.1), is the sum of the price $\frac{\bar{D}_n}{r}$ that would arise if arbitrageurs had infinite wealth, and of a discount that is proportional to the asset's covariance $(\Sigma(s+u))_n$ with aggregate supply $s+u$. Changes in arbitrageur wealth affect the discount, and hence their effect is proportional to the covariance. Since the asset's expected return, given in (6.2), is proportional to the same covariance, aggregate illiquidity is a priced risk factor and explains expected returns perfectly, as in the case of zero supply. The equivalence between positive and zero supply does not hold for short-lived hedgers because the proportionality coefficients (e.g., the function $g(w_t)$) depend on both s and u rather than only on their sum.

7 Conclusion

We develop a dynamic model of liquidity provision, in which hedgers can trade multiple risky assets with arbitrageurs. Arbitrageurs have CRRA utility over consumption, and their capital matters because of wealth effects. We strip out frictions such as asymmetric information and borrowing constraints. At the same time, we depart from most frictionless asset-pricing models by fixing the riskless rate and by suppressing wealth effects for the arbitrageurs' counterparties. Under this combination of assumptions, we prove general analytical results on equilibrium prices and allocations. We characterize, in particular, how arbitrageurs' risk aversion, the endogenous risk that they generate, and the pricing of that risk, depend on arbitrageur wealth and hedger characteristics. We also show that arbitrageur wealth is the single priced risk factor, and that aggregate illiquidity, which declines in wealth, captures that factor.

One important extension of our model is to allow the supply u coming from hedgers to be time-varying and stochastic. Such an extension would give our measure of illiquidity (5.1) stronger empirical content because that measure is based on supply shocks. We could also compare our measure to measures commonly used in empirical work, e.g., Amihud (2002) and Pastor and Stambaugh (2003), and identify their properties when volume arises both because of supply shocks affecting liquidity demanders and wealth effects affecting liquidity providers. In a similar spirit, supply shocks may generate a tighter relationship between volatility and our measure of illiquidity. Indeed, volatility in our model is driven by wealth effects of arbitrageurs and is hump-shaped in wealth, while illiquidity is defined based on supply shocks affecting hedgers and is decreasing in wealth.

Extending our model to stochastic u could strengthen our interpretation of arbitrageurs as specialized liquidity providers. Indeed, a common view of liquidity providers (e.g., Grossman and Miller (1988)) is that they absorb temporary imbalances between demand and supply. A natural interpretation of these imbalances within our model is as shocks to u. In the presence of these shocks, liquidity provision becomes distinct from sharing the aggregate risk in the economy, which also includes the supply s coming from issuers. To model this idea, we would need to introduce additional agents to the model who absorb part of s but are unable to identify shocks to u or trade on them. Under such an extension, the result of Section 6 that s and u are symmetric in terms of their effects on prices and expected returns would break down. Arbitrageur wealth or variables related to it, such as illiquidity, may remain the single priced risk factor, however.

References

Acharya, Viral, Lars Lochstoer, and Tarun Ramadorai, 2013, Limits to arbitrage and hedging: Evidence from commodity markets, *Journal of Financial Economics* 109, 441–465.

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

Adrian, Tobias, Erkko Etula, and Tyler Muir, 2014, Financial intermediaries and the cross-section of asset returns, *Journal of Finance* 69, 2557–2596.

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

Aragon, George, and Philip Strahan, 2012, Hedge funds as liquidity providers: Evidence from the Lehman bankruptcy, *Journal of Financial Economics* 103, 570587.

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

Bhamra, Harjoat, 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.

Bogachev, Vladimir, Nikolai Krylov, Michael Röckner, and Stanislav Shaposhnikov, 2015, *Fokker-Planck-Kolmogorov Equations* (American Mathematical Society, Providence).

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

Brunnermeier, Markus, and Yuliy Sannikov, 2014, A macroeconomic model with a financial sector, *American Economic Review* 104, 379–421.

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

Chan, 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.

Cheng, Ing-Haw, and Wei Xiong, 2014, Why do hedgers trade so much, *Journal of Legal Studies* 43, S183–S207.

Cheng, Ing-Haw, and Wei Xiong, 2015, Convective risk flows in commodity futures markets, *Review of Finance* 19, 1733–1781.

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.

Coval, Joshua, and Erik Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 475–512.

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

Ehling, Paul, and Christian Heyerdahl-Larsen, 2017, Correlations, *Management Science* 63, 1919–1937.

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

Garleanu, Nicolae, and Stavros Panageas, 2015, Young, old, conservative and bold: The implications of heterogeneity and finite lives for asset pricing, *Journal of Political Economy* 123, 670–685.

Garleanu, Nicolae, and Lasse 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.

Gromb, Denis, and Dimitri Vayanos, 2017, The dynamics of financially constrained arbitrage, *Journal of Finance* forthcoming.

Grossman, Sanford, and Merton Miller, 1988, Liquidity and market structure, *Journal of Finance* 43, 617–637.

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

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

He, Zhiguo, Bryan Kelly, and Asaf Manela, 2017, Intermediary asset pricing: New evidence from many asset classes, *Journal of Financial Economics* 126, 1–35.

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.

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

Isaenko, Sergey, 2008, Equilibrium theory of stock market crashes, working paper Concordia University.

Johnson, Timothy, 2008, Volume, liquidity, and liquidity risk, *Journal of Financial Economics* 87, 388–417.

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

Kang, Wenjin, Geert Rouwenhorst, and Ke Tang, 2016, A tale of two premia: The role of hedgers and speculators in commodity futures markets, working paper Yale School of Management.

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.

Liu, Qi, Lei Lu, Bo Sun, and Hongjun Yan, 2017, A model of anomany discovery, working paper DePaul University.

Longstaff, Francis, and Jiang Wang, 2012, Asset pricing and the credit market, *Review of Financial Studies* 25, 3169–3215.

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

Murray, Francis J, and Kenneth S Miller, 2013, *Existence theorems for ordinary differential equations* (Courier Corporation).

Pastor, Lubos, and Robert 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.

Vayanos, Dimitri, and Jiang Wang, 2013, Market liquidity—theory and empirical evidence, in George Constantinides, Milton Harris, and Rene Stulz, eds.: *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.

APPENDIX

A Proofs of the Results in Section 3

Proof of Proposition 3.1: The Bellman equation is

$$\bar{\rho}V = \max_{\bar{c}_t, x_t} \left\{ u(\bar{c}_t) + V_v \mu_{vt} + \frac{1}{2} V_{vv} \sigma_{vt}^\top \sigma_{vt} + V_w \mu_{wt} + \frac{1}{2} V_{ww} \sigma_{wt}^\top \sigma_{wt} + V_{vw} \sigma_{vt}^\top \sigma_{wt} \right\}, \tag{A.1}$$

where $u(\bar{c}_t) = e^{-\frac{\alpha}{r}\bar{c}_t}$, (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. Substituting dD_t from (2.1) into (3.3) and (3.8), we find that the drifts and diffusions are

$$\mu_{vt} = rv_t - \bar{c}_t + x_t^\top (\bar{D} - \pi_t) + u^\top \bar{D}, \tag{A.2}$$

$$\sigma_{vt} = \sigma(x_t + u), \tag{A.3}$$

for the hedger, and

$$\mu_{wt} = rw_t - c_t + y_t^\top (\bar{D} - \pi_t), \tag{A.4}$$

$$\sigma_{wt} = \sigma y_t, \tag{A.5}$$

for arbitrageurs.

Substituting (3.4) and (A.2)-(A.5) into (A.1), we can write the latter equation as

$$\begin{aligned} - \bar{\rho} e^{-[\alpha v_t + F(w_t)]} &= \max_{\bar{c}_t, x_t} \Big\{ -e^{-\frac{\alpha}{r}\bar{c}_t} + \alpha e^{-[\alpha v_t + F(w_t)]} \left[rv_t - \bar{c}_t + x_t^\top (\bar{D} - \pi_t) + u^\top \bar{D} \right] \\ & - \frac{1}{2}\alpha^2 e^{-[\alpha v_t + F(w_t)]} (x_t + u)^\top \Sigma (x_t + u) + F'(w_t) e^{-[\alpha v_t + F(w_t)]} \left[rw_t - c_t + y_t^\top (\bar{D} - \pi_t) \right] \\ & + \frac{1}{2} \left[F''(w_t) - F'(w_t)^2 \right] e^{-[\alpha v_t + F(w_t)]} y_t^\top \Sigma y_t - \alpha F'(w_t) e^{-[\alpha v_t + F(w_t)]} (x_t + u)^\top \Sigma y_t \Big\}. \end{aligned} \tag{A.6}$$

The first-order conditions with respect to $\bar{c}_t$ and x_t yield (3.5) and (3.6), respectively. ∎

Proof of Proposition 3.2: We can write the hedger's maximization problem as

$$\max_{x_t} \left\{ \mu_{vt} - \frac{\alpha}{2} \sigma_{vt}^\top \sigma_{vt} \right\}, \tag{A.7}$$

where μ_{vt} is given by (A.2) with $\bar{c}_t = 0$, and σ_{vt} is given by (A.3). Using (A.2) and (A.3), we find that the problem (A.7) is equivalent to

$$\max_{x_t} \left\{ x_t^\top (\bar{D} - \pi_t) - \frac{\alpha}{2}(x_t + u)^\top \Sigma (x_t + u) \right\}. \tag{A.8}$$

The first-order condition with respect to x_t yields (3.7). ∎

Proof of Proposition 3.3: The Bellman equation is

$$\rho V = \max_{\hat{c}_t, \hat{y}_t} \left\{ u(\hat{c}_t) + V_{\hat{w}} \mu_{\hat{w}t} + \frac{1}{2} V_{\hat{w}\hat{w}} \sigma_{\hat{w}t}^\top \sigma_{\hat{w}t} + V_w \mu_{wt} + \frac{1}{2} V_{ww} \sigma_{wt}^\top \sigma_{wt} + V_{\hat{w}w} \sigma_{\hat{w}t}^\top \sigma_{wt} \right\}, \tag{A.9}$$

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' aggregate wealth. The quantities (μ_{wt}, σ_{wt}) are given by (A.4) and (A.5), respectively. The quantities $(\mu_{\hat{w}t}, \sigma_{\hat{w}t})$ are given by the counterparts of these equations with hat signs, i.e.,

$$\mu_{\hat{w}t} = r\hat{w}_t - \hat{c}_t + \hat{y}_t^\top (\bar{D} - \pi_t), \tag{A.10}$$

$$\sigma_{\hat{w}t} = \sigma \hat{y}_t. \tag{A.11}$$

When $\gamma \neq 1$, we substitute (3.9), (A.4), (A.5), (A.10) and (A.11) into (A.9) to write it as

$$\begin{aligned} \rho q(w_t) \frac{\hat{w}_t^{1-\gamma}}{1-\gamma} = \max_{\hat{c}_t, \hat{y}_t} \Bigg\{ & \frac{\hat{c}_t^{1-\gamma}}{1-\gamma} + q(w_t) \hat{w}_t^{-\gamma} \left(r\hat{w}_t - \hat{c}_t + \hat{y}_t^\top (\bar{D} - \pi_t) \right) - \frac{1}{2} q(w_t) \gamma \hat{w}_t^{-\gamma-1} \hat{y}_t^\top \Sigma \hat{y}_t \\ & + q'(w_t) \frac{\hat{w}_t^{1-\gamma}}{1-\gamma} \left(r w_t - c_t + y_t^\top (\bar{D} - \pi_t) \right) + \frac{1}{2} q''(w_t) \frac{\hat{w}_t^{1-\gamma}}{1-\gamma} y_t^\top \Sigma y_t + q'(w_t) \hat{w}_t^{-\gamma} \hat{y}_t^\top \Sigma y_t \Bigg\}. \end{aligned} \tag{A.12}$$

The first-order conditions with respect to $\hat{c}_t$ and $\hat{y}_t$ yield (3.11) and (3.12), respectively. When $\gamma = 1$, we substitute (3.10), (A.4), (A.5), (A.10) and (A.11) into (A.9) to write it as

$$\begin{aligned} \rho \left(\frac{1}{\rho} \log(\hat{w}_t) + q_1(w_t) \right) = \max_{\hat{c}_t, \hat{y}_t} \Bigg\{ & \log(\hat{c}_t) + \frac{1}{\rho \hat{w}_t} \left(r\hat{w}_t - \hat{c}_t + \hat{y}_t^\top (\bar{D} - \pi_t) \right) - \frac{1}{2\rho \hat{w}_t^2} \hat{y}_t^\top \Sigma \hat{y}_t \\ & + q_1'(w_t) \left(r w_t - c_t + y_t^\top (\bar{D} - \pi_t) \right) + \frac{1}{2} q_1''(w_t) y_t^\top \Sigma y_t \Bigg\}. \end{aligned} \tag{A.13}$$

The first-order conditions with respect to $\hat{c}_t$ and $\hat{y}_t$ yield (3.11) and (3.12), respectively, for $q(w_t) = \frac{1}{\rho}$. ∎

Proof of Proposition 3.4: Since in equilibrium $\hat{c}_t = c_t$ and $\hat{w}_t = w_t$, (3.11) implies that

$$c_t = q(w_t)^{-\frac{1}{\gamma}} w_t. \tag{A.14}$$

Substituting (3.17) and (3.18) into (3.12) and using the definition of $A(w_t)$ in (3.14), we find

$$\hat{y}_t = \frac{\alpha \hat{w}_t}{[\alpha + A(w_t) - F'(w_t)]w_t} u. \tag{A.15}$$

When $\gamma \neq 1$, we substitute (3.11), (3.17), (3.18), (A.14), and (A.15) into the arbitrageur's Bellman equation (A.12). The terms in $\hat{w}_t$ cancel, and the resulting equation is

$$\begin{aligned}\rho q(w_t) =& q(w_t)^{1-\frac{1}{\gamma}} + \left(q'(w_t) + \frac{(1-\gamma)q(w_t)}{w_t}\right)\left(rw_t - q(w_t)^{-\frac{1}{\gamma}} w_t + \frac{\alpha^2 u^\top \Sigma u A(w_t)}{[\alpha + A(w_t) - F'(w_t)]^2}\right) \\ &+ \frac{1}{2}\left(q''(w_t) - \frac{\gamma(1-\gamma)q(w_t)}{w_t^2} + \frac{2(1-\gamma)q'(w_t)}{w_t}\right)\frac{\alpha^2 u^\top \Sigma u}{[\alpha + A(w_t) - F'(w_t)]^2}.\end{aligned} \tag{A.16}$$

Dividing both sides by $q(w_t)$ and rearranging, we find

$$\begin{aligned}\rho =& q(w_t)^{-\frac{1}{\gamma}} + \left(\frac{q'(w_t)}{q(w_t)} + \frac{1-\gamma}{w_t}\right)\left(rw_t - q(w_t)^{-\frac{1}{\gamma}} w_t\right) \\ &+ \frac{1}{2}\left(\frac{q''(w_t)}{q(w_t)} - \frac{\gamma(1-\gamma)}{w_t^2} + \frac{2(1-\gamma)q'(w_t)}{q(w_t)w_t} + 2\left(\frac{q'(w_t)}{q(w_t)} + \frac{1-\gamma}{w_t}\right)A(w_t)\right)\frac{\alpha^2 u^\top \Sigma u}{[\alpha + A(w_t) - F'(w_t)]^2} \\ =& r + \left(q(w_t)^{-\frac{1}{\gamma}} - r\right)A(w_t)w_t \\ &+ \frac{1}{2}\left(\frac{q''(w_t)}{q(w_t)} + \frac{\gamma(1+\gamma)}{w_t^2} - \frac{2\gamma q'(w_t)}{q(w_t)w_t} - 2A(w_t)^2\right)\frac{\alpha^2 u^\top \Sigma u}{[\alpha + A(w_t) - F'(w_t)]^2},\end{aligned} \tag{A.17}$$

where the last step follows from the definition of $A(w_t)$ in (3.14). Differentiating (3.14), we find

$$\begin{aligned}A'(w_t) =& -\frac{\gamma}{w_t^2} - \frac{q''(w_t)}{q(w_t)} + \frac{q'(w_t)^2}{q(w_t)^2} \\ =& -\frac{\gamma(1+\gamma)}{w_t^2} - \frac{q''(w_t)}{q(w_t)} + \frac{2\gamma q'(w_t)}{q(w_t)w_t} + A(w_t)^2,\end{aligned} \tag{A.18}$$

where the second step follows again from (3.14). Substituting (A.18) into (A.17), we find

$$\rho = r + \left(q(w_t)^{-\frac{1}{\gamma}} - r\right)A(w_t)w_t - \frac{1}{2}\left(A'(w_t) + A(w_t)^2\right)\frac{\alpha^2 u^\top \Sigma u}{[\alpha + A(w_t) - F'(w_t)]^2}. \tag{A.19}$$

Rearranging (A.19) and using the definition of z in (3.20), we find the ODE (3.21).

When $\gamma = 1$, $q(w_t) = \frac{1}{\rho}$ and $A(w_t) = \frac{1}{w_t}$. The ODE (3.21) holds for these values. The arbitrageur's Bellman equation (A.13) yields an ODE involving $q_1(w_t)$. To derive it we substitute (3.11), (3.17), (3.18), (A.14), (A.15), $q(w_t) = \frac{1}{\rho}$ and $A(w_t) = \frac{1}{w_t}$ into (A.13). The terms in $\hat{w}_t$ cancel, and the resulting equation is

$$\begin{aligned}\rho q_1(w_t) =& \log(\rho) + \left(q_1'(w_t) + \frac{1}{\rho w_t}\right)\left(rw_t - \rho w_t + \frac{\alpha^2 u^\top \Sigma u \frac{1}{w_t}}{\left[\alpha + \frac{1}{w_t} - F'(w_t)\right]^2}\right) \\ &+ \frac{1}{2}\left(q_1''(w_t) - \frac{1}{\rho w_t^2}\right)\frac{\alpha^2 u^\top \Sigma u}{\left[\alpha + \frac{1}{w_t} - F'(w_t)\right]^2}. \end{aligned} \tag{A.20}$$

Using (3.5) and (3.6), we can write the terms in the first line of the hedger's Bellman equation (A.6) as

$$\begin{aligned}& - e^{-\frac{\alpha}{r}\bar{c}_t} + \alpha e^{-[\alpha v_t + F(w_t)]}\left[rv_t - \bar{c}_t + x_t^\top(\bar{D} - \pi_t) + u^\top \bar{D}\right] \\ &= -re^{-[\alpha v_t + F(w_t)]} + \alpha e^{-[\alpha v_t + F(w_t)]}\left[\frac{r}{\alpha}\left[\log(r) - F(w_t)\right] + u^\top \bar{D} + \alpha x_t^\top \Sigma(x_t + u) + F'(w_t) x_t^\top \Sigma y_t\right].\end{aligned}$$

Substituting into (A.6), we can write that equation as

$$\begin{aligned}\bar{\rho} =& r + rF(w_t) - r\log(r) - \alpha u^\top \bar{D} - \frac{1}{2}\alpha^2(x_t - u)^\top \Sigma(x_t + u) - F'(w_t)\left[rw_t - c_t + y_t^\top(\bar{D} - \pi_t)\right] \\ &- \frac{1}{2}\left[F''(w_t) - F'(w_t)^2\right] y_t^\top \Sigma y_t + \alpha F'(w_t) u^\top \Sigma y_t. \end{aligned} \tag{A.21}$$

The last four terms in the right-hand side of (A.21) are equal to

$$\begin{aligned}& - \frac{1}{2}\alpha^2 x_t^\top \Sigma x_t + \frac{1}{2}\alpha^2 u^\top \Sigma u - F'(w_t)\left[rw_t - c_t + A(w_t) y_t^\top \Sigma y_t\right] \\ &- \frac{1}{2}\left[F''(w_t) - F'(w_t)^2\right] y_t^\top \Sigma y_t + \alpha F'(w_t) u^\top \Sigma y_t \\ &= -\frac{1}{2}\alpha^2 y_t^\top \Sigma y_t + \frac{1}{2}\alpha^2 u^\top \Sigma u - F'(w_t)\left[rw_t - c_t + A(w_t) y_t^\top \Sigma y_t\right] \\ &- \frac{1}{2}\left[F''(w_t) - F'(w_t)^2\right] y_t^\top \Sigma y_t + \alpha F'(w_t) u^\top \Sigma y_t \\ &= -F'(w_t)(rw_t - c_t) - \frac{\alpha^2 u^\top \Sigma u}{2\left[\alpha + A(w_t) - F'(w_t)\right]^2} \\ &\times \left[\alpha^2 - \left[\alpha + A(w_t) - F'(w_t)\right]^2 + 2F'(w_t)A(w_t) + F''(w_t) - F'(w_t)^2 - 2F'(w_t)\left[\alpha + A(w_t) - F'(w_t)\right]\right] \\ &= -F'(w_t)(rw_t - c_t) - \frac{\alpha^2 u^\top \Sigma u \left\{F''(w_t) - A(w_t)\left[2\alpha + A(w_t) - 2F'(w_t)\right]\right\}}{2\left[\alpha + A(w_t) - F'(w_t)\right]^2},\end{aligned}$$

where the first step follows from (3.13), the second from (3.15), and the third from (3.18). Substituting into (A.21) and simplifying, we find

$$\begin{aligned}\bar{\rho} =& r + rF(w_t) - r\log(r) - \alpha u^\top \bar{D} - F'(w_t)(rw_t - c_t) \\ &- \frac{\alpha^2 u^\top \Sigma u \left\{F''(w_t) - A(w_t)\left[2\alpha + A(w_t) - 2F'(w_t)\right]\right\}}{2\left[\alpha + A(w_t) - F'(w_t)\right]^2}.\end{aligned} \tag{A.22}$$

Substituting c_t from (A.14) and rearranging, we find (3.22). ∎

Proof of Theorem 3.1: We start with the case where hedgers are short-lived. The second-order ODE (3.21) for $q(w_t)$ is equivalent to the system of two first-order ODEs for $(q(w_t), A(w_t))$:

$$q'(w_t) = q(w_t)\left(\frac{\gamma}{w_t} - A(w_t)\right), \tag{A.23}$$

$$A'(w_t) = -A(w_t)^2 + \frac{\left[\alpha + A(w_t)\right]^2}{z}\left(\frac{q(w_t)^{-\frac{1}{\gamma}} - r}{\rho - r}A(w_t)w_t - 1\right), \tag{A.24}$$

which follow from rearranging (3.14) and (3.21), respectively. Using the functions $Q(w_t) \equiv q(w_t)^{-\frac{1}{\gamma}}$ and $R(w_t) \equiv A(w_t)w_t$, we can convert the system of (A.23) and (A.24) into one for $(Q(w_t), R(w_t))$:

$$Q'(w_t) = \frac{Q(w_t)\left[R(w_t) - \gamma\right]}{\gamma w_t}, \tag{A.25}$$

$$R'(w_t)w_t = R(w_t)\left[1 - R(w_t)\right] + \frac{\left[\alpha w_t + R(w_t)\right]^2}{z}\left(\frac{Q(w_t) - r}{\rho - r}R(w_t) - 1\right). \tag{A.26}$$

Since the function

$$(w_t, Q, R) \longrightarrow \begin{pmatrix} \frac{Q(R-\gamma)}{w_t} \\ \frac{R(1-R)}{w_t} + \frac{(\alpha w_t + R)^2}{z w_t}\left(\frac{Q-r}{\rho-r}R - 1\right) \end{pmatrix}$$

is continuously differentiable for $(w, Q, R) \in (0,\infty) \times (-\infty,\infty) \times (-\infty,\infty)$, it is locally Lipschitz in that set. Hence, for any $\epsilon > 0$, the system of (A.25) and (A.26) has a unique solution in a neighborhood of ϵ with initial conditions $Q(\epsilon) > 0$ and $R(\epsilon) > 0$. We can extend that solution maximally to the left and to the right of ϵ, over an interval $\mathcal{I}$. That solution satisfies $Q(w_t) > 0$ for all $w_t \in \mathcal{I}$ because of (A.25). We next derive properties of solutions to the system of (A.25) and (A.26), as well as the existence result, through a number of lemmas.

Lemma A.1 (Limits at zero and infinity) *Consider a solution $Q(w_t) > 0$ and $R(w_t)$ to the system of (A.25) and (A.26), defined over the interval $(0, \infty)$.*

- *If* $\lim_{w_t \to 0} R(w_t) \in (0, \infty)$, *then this limit is equal to* K *if* $\gamma < K$ *and to* γ *if* $\gamma > K$. *In the former case* $\lim_{w_t \to 0} Q(w_t) = 0$ *and in the latter case* $\lim_{w_t \to 0} Q(w_t) \in (0, \infty)$.

- *If* $\lim_{w_t \to \infty} Q(w_t) \in (0, \infty)$, *then this limit is equal to* $r + \frac{\rho - r}{\gamma}$, *and* $\lim_{w_t \to \infty} R(w_t) = \gamma$.

Proof: To derive the limits at zero, we start by observing that the finiteness of $\lim_{w_t \to 0} R(w_t)$ implies that if $\lim_{w_t \to 0} R'(w_t) w_t$ exists then it must be zero. Indeed, if $\lim_{w_t \to 0} R'(w_t) w_t \neq 0$, then $|R'(w_t)| \geq \frac{\ell}{w_t}$ for $w_t < \epsilon$ and for positive ϵ and ℓ. Since, however, for $w_t < \epsilon$

$$R(w_t) = R(\epsilon) + \int_{\epsilon}^{w_t} R'(\hat{w}_t) d\hat{w}_t \Rightarrow |R(w_t) - R(\epsilon)| \geq \int_{w_t}^{\epsilon} \frac{\ell}{\hat{w}_t} d\hat{w}_t = \ell \log\left(\frac{\epsilon}{w_t}\right),$$

$\lim_{w_t \to 0} R(w_t)$ would be plus or minus infinity, a contradiction.

If $\lim_{w_t \to 0} R(w_t) < \gamma$, then (A.25) implies $\lim_{w_t \to 0} Q(w_t) = \infty$. The latter equation, together with (A.26) and $\lim_{w_t \to 0} R(w_t) \neq 0$, imply $\lim_{w_t \to 0} R'(w_t) w_t = \infty$, a contradiction. If $\lim_{w_t \to 0} R(w_t) > \gamma$, then (A.25) implies $\lim_{w_t \to 0} Q(w_t) = 0$ and (A.26) implies

$$\lim_{w_t \to 0} R'(w_t) w_t = \lim_{w_t \to 0} R(w_t) G\left(\lim_{w_t \to 0} R(w_t)\right). \tag{A.27}$$

Setting $\lim_{w_t \to 0} R'(w_t) w_t = 0$ in (A.27), and using $\lim_{w_t \to 0} R(w_t) \neq 0$, we find $G\left(\lim_{w_t \to 0} R(w_t)\right) = 0$. Since the quadratic polynomial $G(\gamma)$ defined in (3.23) has the unique positive root K, $\lim_{w_t \to 0} R(w_t) = K$. Suppose, finally, that $\lim_{w_t \to 0} R(w_t) = \gamma$ and $\lim_{w_t \to 0} R'(w_t) w_t$ exists. Setting $\lim_{w_t \to 0} R'(w_t) w_t = 0$ in (A.26), we find

$$\begin{aligned}
&\gamma(1-\gamma) + \frac{\gamma^2}{z}\left(\frac{\lim_{w_t \to 0} Q(w_t) - r}{\rho - r}\gamma + 1\right) = 0 \\
&\Leftrightarrow 1 - \gamma + \frac{\gamma}{z}\left(\frac{\lim_{w_t \to 0} Q(w_t) - r}{\rho - r}\gamma + 1\right) = 0. \qquad \text{(A.28)}
\end{aligned}$$

Since K is the unique positive root of $G(\gamma)$ and the quadratic term of $G(\gamma)$ is negative, (A.28) has a positive solution for $\lim_{w_t \to 0} Q(w_t)$ if $\gamma > K$.

If $\gamma > K$, then the case $\lim_{w_t \to 0} R(w_t) > \gamma$ is not possible because it would imply $\lim_{w_t \to 0} R(w_t) = K < \gamma$. Hence, $\lim_{w_t \to 0} R(w_t) = \gamma$ and $\lim_{w_t \to 0} Q(w_t) \in (0, \infty)$. If $\gamma < K$, then the case $\lim_{w_t \to 0} R(w_t) = \gamma$ is not possible because it would imply $\lim_{w_t \to 0} Q(w_t) < 0$. Hence, $\lim_{w_t \to 0} R(w_t) = K$ and $\lim_{w_t \to 0} Q(w_t) = 0$.

To complete the proof for the limits at zero, we need to show that in the case $\lim_{w_t \to 0} R(w_t) = \gamma$, $\lim_{w_t \to 0} R'(w_t) w_t$ exists. From (A.26), this is equivalent to showing that $\lim_{w_t \to 0} Q(w_t)$ exists. We proceed by contradiction and assume that $\lim_{w_t \to 0} Q(w_t)$ does not exist, and hence $\limsup_{w_t \to 0} Q(w_t) > \liminf_{w_t \to 0} Q(w_t)$. This means that as w_t goes to zero $Q(w_t)$ oscillates more and more rapidly between maxima and minima whose distance is bounded away from zero. Differentiating (A.25) at a maximizer and at a minimizer of $Q(w_t)$, we find

$$\begin{aligned} Q''(w_t) &= Q'(w_t) \frac{R(w_t) - \gamma}{\gamma w_t} + Q(w_t) \left(\frac{R'(w_t)}{\gamma w_t} - \frac{R(w_t) - \gamma}{\gamma w_t^2} \right) \\ &= \frac{Q(w_t) R'(w_t)}{\gamma w_t} \\ &= \frac{Q(w_t) \left[\gamma(1-\gamma) + \frac{(\alpha w_t + \gamma)^2}{z} \left(\frac{Q(w_t) - r}{\rho - r} \gamma - 1 \right) \right]}{\gamma w_t^2}, \end{aligned} \tag{A.29}$$

where the second step follows from $R(w_t) = \gamma$, which holds at a maximizer and a minimizer of $Q(w_t)$ because of (A.25), and the third step follows from (A.26). Since $Q''(w_t) < 0$ at a maximizer, (A.29) implies

$$\gamma(1-\gamma) + \frac{(\alpha w_t + \gamma)^2}{z} \left(\frac{Q(w_t) - r}{\rho - r} \gamma - 1 \right) < 0.$$

Taking the limit along a sequence of maximizers as w_t goes to zero, we find

$$\gamma(1-\gamma) + \frac{\gamma^2}{z} \left(\frac{\limsup_{w_t \to 0} Q(w_t) - r}{\rho - r} \gamma - 1 \right) \le 0. \tag{A.30}$$

Since $Q''(w_t) > 0$ at a minimizer, (A.29) likewise implies that

$$\gamma(1-\gamma) + \frac{\gamma^2}{z} \left(\frac{\liminf_{w_t \to 0} Q(w_t) - r}{\rho - r} \gamma - 1 \right) \ge 0. \tag{A.31}$$

Equations (A.30) and (A.31) imply that $\limsup_{w_t \to 0} Q(w_t) \le \liminf_{w_t \to 0} Q(w_t)$, a contradiction.

Consider next the limits at infinity. If $\lim_{w_t \to \infty} R(w_t)$ exists, then it has to equal γ, otherwise (A.25) would imply that $\lim_{w_t \to \infty} Q(w_t)$ is either zero or infinity. Moreover,

$$\frac{\lim_{w_t \to \infty} Q(w_t) - r}{\rho - r} \gamma - 1 = 0 \Rightarrow \lim_{w_t \to \infty} Q(w_t) = r + \frac{\rho - r}{\gamma},$$

otherwise (A.26) would imply that $\lim_{w_t \to 0} R'(w_t)$ is plus or minus infinity, which contradicts the finiteness of $\lim_{w_t \to 0} R(w_t)$.

To complete the proof for the limits at infinity, we need to show that $\lim_{w_t \to \infty} R(w_t)$ exists. We proceed by contradiction and assume that $\lim_{w_t \to \infty} R(w_t)$ does not exist, and hence $\limsup_{w_t \to \infty} R(w_t) > \liminf_{w_t \to \infty} R(w_t)$. Since $R(w_t)$ oscillates between values close to $\limsup_{w_t \to \infty} R(w_t)$ and values close to $\liminf_{w_t \to \infty} R(w_t)$, there exists $\xi \in (\liminf_{w_t \to \infty} R(w_t), \limsup_{w_t \to \infty} R(w_t))$ and a sequence $\{w_{tn}\}_{n \in \mathbb{N}}$ converging to infinity such that $R(w_{tn}) = \xi$ and $R'(w_{tn})$ alternates between being non-positive and non-negative. Equation (A.26) implies, however, that for large n, the sign of $R'(w_{tn})$ is the same as of

$$\frac{\lim_{w_t \to \infty} Q(w_t) - r}{\rho - r} \xi - 1,$$

which does not change and can be chosen to be non-zero by varying ξ, a contradiction. ∎

Lemma A.2 (Single crossing of solutions) *Consider two solutions* $(Q_1(w_t), R_1(w_t))$ *and* $(Q_2(w_t), R_2(w_t))$ *to the system of (A.25) and (A.26) with initial conditions* $Q_1(\epsilon) > Q_2(\epsilon) > 0$ *and* $R_1(\epsilon) = R_2(\epsilon) > 0$ *for* $\epsilon > 0$. *The solutions compare as follows:*

- $Q_1(w_t) > Q_2(w_t)$ *for all* w_t *for which the solutions are defined.*
- $R_1(w_t) > R_2(w_t)$ *for all* $w_t > \epsilon$, *and* $R_1(w_t) < R_2(w_t)$ *for all* $w_t < \epsilon$.

Proof: We first show the inequalities for $w_t > \epsilon$. Since $Q_1(\epsilon) > Q_2(\epsilon) > 0$, (A.26) implies $R_1'(\epsilon) > R_2'(\epsilon)$. Since, in addition, $R_1(\epsilon) = R_2(\epsilon)$, $R_1(w_t) > R_2(w_t)$ for w_t close to and larger than ϵ. Proceeding by contradiction, suppose that there exists $w_t > \epsilon$ such that $Q_1(w_t) \leq Q_2(w_t)$ or $R_1(w_t) \leq R_2(w_t)$. The infimum m within that set is strictly larger than ϵ since $Q_1(w_t) > Q_2(w_t)$ and $R_1(w_t) > R_2(w_t)$ for w_t close to and larger than ϵ. Since the same inequalities hold for all $w_t \in (\epsilon, m)$, $Q_1(m) \geq Q_2(m)$ and $R_1(m) \geq R_2(m)$, with one of these inequalities being an equality. Integrating (A.25) from ϵ to m, we find

$$\log\left(\frac{Q(m)}{Q(\epsilon)}\right) = \int_\epsilon^m \frac{R(w_t) - \gamma}{\gamma w_t} dw_t. \tag{A.32}$$

Since $R_1(w_t) > R_2(w_t)$ for all $w_t \in (\epsilon, m)$, (A.32) implies $Q_1(m) > Q_2(m)$. Hence, $R_1(m)$ must be equal to $R_2(m)$. Since $R_1(w_t) > R_2(w_t)$ for all $w_t \in (\epsilon, m)$, $R_1'(m) \leq R_2'(m)$. Equation (A.26) then implies $Q_1(m) \leq Q_2(m)$, a contradiction.

The inequalities for $w_t < \epsilon$ follow from a similar argument. Since $R_1(\epsilon) = R_2(\epsilon)$ and $R_1'(\epsilon) > R_2'(\epsilon)$, $R_1(w_t) < R_2(w_t)$ for w_t close to and smaller than ϵ. Proceeding by contradiction, suppose that there exists $w_t < \epsilon$ such that $Q_1(w_t) \leq Q_2(w_t)$ or $R_1(w_t) \geq R_2(w_t)$. The supremum m within that set is strictly smaller than ϵ since $Q_1(w_t) > Q_2(w_t)$ and $R_1(w_t) < R_2(w_t)$ for w_t close to and smaller than ϵ. Since the same inequalities hold for all $w_t \in (m, \epsilon)$, $Q_1(m) \geq Q_2(m)$ and $R_1(m) \leq R_2(m)$, with one of these inequalities being an equality. Since $R_1(w_t) < R_2(w_t)$ for all $w_t \in (m, \epsilon)$, (A.32) implies $Q_1(m) > Q_2(m)$. Hence, $R_1(m)$ must be equal to $R_2(m)$. Since $R_1(w_t) < R_2(w_t)$ for all $w_t \in (m, \epsilon)$, $R_1'(m) \leq R_2'(m)$. Equation (A.26) then implies $Q_1(m) \leq Q_2(m)$, a contradiction. ∎

Lemma A.3 (Existence in finite interval) *For any $\epsilon > 0$ and $M > \epsilon$ large enough, there exists a unique solution to the system of (A.25) and (A.26) that is defined over an interval including $[\epsilon, M]$ and that satisfies $R(\epsilon) = \max\{\gamma, K\}$, $R(M) = \gamma$ and $Q(\epsilon) > 0$.*

Proof: We will consider solutions to the system of (A.25) and (A.26) with $R(\epsilon) = \max\{\gamma, K\}$, and show that by varying $Q(\epsilon)$ we can vary $R(M)$ from negative values to large positive values. We will then use a continuity argument to find a suitable $Q(\epsilon)$, and the single-crossing argument in Lemma A.2 to show that such a $Q(\epsilon)$ is unique.

Consider the solution to the system of (A.25) and (A.26) with initial conditions $R(\epsilon) = \max\{\gamma, K\}$ and $Q(\epsilon) = 0$, and extend it maximally to the right of ϵ, over an interval $\mathcal{I}$. Since $Q(\epsilon) = 0$, (A.25) implies $Q(w_t) = 0$ for all $w_t \in \mathcal{I}$. Equation (A.25) implies that $R'(\epsilon)$ has the same sign as

$$\begin{aligned}
&\max\{\gamma, K\}\left(1 - \max\{\gamma, K\}\right) - \frac{(\alpha w + \max\{\gamma, K\})^2}{z}\left(\frac{r \max\{\gamma, K\}}{\rho - r} + 1\right) \\
&< \max\{\gamma, K\}\left(1 - \max\{\gamma, K\}\right) - \frac{(\max\{\gamma, K\})^2}{z}\left(\frac{r \max\{\gamma, K\}}{\rho - r} + 1\right) \\
&= \max\{\gamma, K\} G\left(\max\{\gamma, K\}\right). \qquad \text{(A.33)}
\end{aligned}$$

If $\gamma \leq K$, then (A.33) is equal to zero because K is a root of $G(\gamma)$, as implied by (3.23). If $\gamma > K$, then (A.33) is negative because K is the unique positive root of $G(\gamma)$ and the quadratic term of $G(\gamma)$ is negative. In both cases, $R'(\epsilon) < 0$ and hence $R(w_t) < \max\{\gamma, K\}$ for w_t close to and larger than ϵ.

We next show that $R(w_t)$ must lie between $\max\{\gamma, K\}$ and the unique negative root of $G(\gamma)$, which we denote by K'. Proceeding by contradiction, suppose that there exists $w_t > \epsilon$ such that

$R(w_t) > \max\{\gamma, K\}$. The infimum m within that set is strictly larger than ϵ since $R(w_t) < \max\{\gamma, K\}$ for w_t close to and larger than ϵ. Since $R(w_t) < \max\{\gamma, K\}$ for all $w_t \in (\epsilon, m)$, $R(m) = \max\{\gamma, K\}$, and hence $R'(m) \geq 0$. Equation (A.25), however, implies that $R'(m)$ has the same sign as

$$\max\{\gamma, K\} (1 - \max\{\gamma, K\}) - \frac{(\alpha m + \max\{\gamma, K\})^2}{z} \left(\frac{r \max\{\gamma, K\}}{\rho - r} + 1 \right),$$

which is negative from the previous argument, a contradiction. Hence, $R(w_t) \leq \{\gamma, K\}$ for all $w_t \in \mathcal{I}$. To show that $R(w_t) \geq K'$ for all $w_t \in \mathcal{I}$, we similarly proceed by contradiction and suppose that there exists $w_t > \epsilon$ such that $R(w_t) < K'$. The infimum m within that set is strictly larger than ϵ, and satisfies $R(m) = K'$ and $R'(m) \leq 0$. Equation (A.25), however, implies that $R'(m)$ has the same sign as

$$\begin{aligned}
&K' \left(1 - K'\right) - \frac{(\alpha m + K')^2}{z} \left(\frac{rK'}{\rho - r} + 1 \right) \\
&> K' \left(1 - K'\right) - \frac{K'^2}{z} \left(\frac{rK'}{\rho - r} + 1 \right) \\
&= K' G\left(K'\right) = 0,
\end{aligned}$$

where the second step follows because $G(K') = 0$ implies $\frac{rK'}{\rho - r} + 1 = \frac{z(K'-1)}{K'} > 0$, and the fourth step follows from K' being a root of $G(\gamma)$. Hence $R'(m) > 0$, a contradiction.

Since $R(w_t)$ lies within a bounded interval, the solution to the system of (A.25) and (A.26) with initial conditions $R(\epsilon) = \max\{\gamma, K\}$ and $Q(\epsilon) = 0$ can be extended to infinity, i.e., $\mathcal{I} = [\epsilon, \infty)$. Since, $R(w_t)$ is bounded and $Q(w_t) = 0$ for all $w_t \in \mathcal{I}$,

$$\frac{r \lim_{w_t \to \infty} R(w_t)}{\rho - r} + 1 = 0 \Rightarrow \lim_{w_t \to \infty} R(w_t) = \frac{r - \rho}{r} < 0,$$

otherwise (A.26) would imply that $\lim_{w_t \to \infty} R'(w_t)$ is plus or minus infinity, which contradicts the boundedness of $R(w_t)$.

Consider next the solution to the system of (A.25) and (A.26) with initial conditions $R(\epsilon) = \max\{\gamma, K\}$ and $Q(\epsilon) = r + \frac{(1+z)(\rho - r)}{\max\{\gamma, K\}}$, and extend it maximally to the right of ϵ, over an interval

$\mathcal{I}$. Equation (A.25) implies that $R'(\epsilon)$ has the same sign as

$$\begin{aligned}
&\max\{\gamma, K\}\left(1-\max\{\gamma, K\}\right)+\frac{\left(\alpha w+\max\{\gamma, K\}\right)^2}{z}\left(\frac{Q(\epsilon)-r}{\rho-r}\max\{\gamma, K\}-1\right)\\
&=\max\{\gamma, K\}\left(1-\max\{\gamma, K\}\right)+\left(\alpha w+\max\{\gamma, K\}\right)^2\\
&>\max\{\gamma, K\}\left(1-\max\{\gamma, K\}\right)+\left(\max\{\gamma, K\}\right)^2=\max\{\gamma, K\}>0,
\end{aligned} \tag{A.34}$$

where the second step follows from the definition of $Q(\epsilon)$. Hence, $R'(\epsilon)>0$. We next show that $R'(w_t)>0$ for all $w_t \in \mathcal{I}$. Proceeding by contradiction, suppose that there exists $w_t>\epsilon$ such that $R'(w_t)\leq 0$. The infimum m within that set is strictly larger than ϵ, and satisfies $R'(m)=0$. Since $R'(w_t)>0$ for all $w_t\in(\epsilon, m)$, $R(w_t)>\max\{\gamma, K\}$ and hence (A.25) implies $Q(m)>Q(\epsilon)$. Equation (A.25) implies that $R'(m)$ has the same sign as

$$\begin{aligned}
&R(m)[1-R(m)]+\frac{[\alpha m+R(m)]^2}{z}\left(\frac{Q(m)-r}{\rho-r}R(m)-1\right)\\
&>R(m)[1-R(m)]+\frac{[\alpha m+R(m)]^2}{z}\left(\frac{Q(\epsilon)-r}{\rho-r}R(m)-1\right)\\
&=R(m)[1-R(m)]+\frac{[\alpha m+R(m)]^2}{z}\left(\frac{(1+z)R(m)}{\max\{\gamma, K\}}-1\right)\\
&>R(m)[1-R(m)]+[\alpha m+R(m)]^2\\
&>R(m)[1-R(m)]+R(m)^2=R(m)>0,
\end{aligned} \tag{A.35}$$

where the second step follows from $Q(m)>Q(\epsilon)$, the third step from the definition of $Q(\epsilon)$, and the fourth step from $R(m)>\max\{\gamma, K\}$. Hence, $R'(m)>0$, a contradiction. Since $R'(w_t)>0$ for all $w_t\in\mathcal{I}$, $R(w_t)$ increases to a limit, which must be infinite. Indeed, if the limit were finite, $\lim_{w_t\to\infty} R'(w_t)=0$, in which case (A.26) would imply $\lim_{w_t\to\infty} Q(w_t)\in(0,\infty)$. This is a contradiction: because $\lim_{w_t\to\infty} R(w_t)>\gamma$, (A.25) implies $\lim_{w_t\to\infty} Q(w_t)=\infty$. Note that $R(w_t)$ can reach its infinite limit when w_t goes to a finite value m, in which case the interval $\mathcal{I}$ is bounded.

We denote by $\Gamma(Q)$ the value of $R(M)$ for the solution to the system of (A.25) and (A.26) with initial conditions $R(\epsilon)=\max\{\gamma, K\}$ and $Q(\epsilon)=Q$. If $\lim_{w_t\to m} R(w_t)=\infty$ for $m\leq M$, then we set $R(M)=\infty$. For large enough M, the function $\Gamma(Q)$ satisfies $\Gamma(0)<0$ because for the solution with initial conditions $R(\epsilon)=\max\{\gamma, K\}$ and $Q(\epsilon)=0$, $\lim_{w_t\to\infty} R(w_t)=\frac{r-\rho}{r}<0$. The function $\Gamma(Q)$ also satisfies $\Gamma\left(r+\frac{(1+z)(\rho-r)}{\max\{\gamma, K\}}\right)>\gamma$ because for the solution with initial conditions $R(\epsilon)=\max\{\gamma, K\}$ and $Q(\epsilon)=r+\frac{(1+z)(\rho-r)}{\max\{\gamma, K\}}$, $\lim_{w_t\to m} R(w_t)=\infty$, where $m>\epsilon$ is the upper end of the interval $\mathcal{I}$. Lemma A.2 implies that for $Q_1 > Q_2$, $\Gamma(Q_1) > \Gamma(Q_2)$ if $\Gamma(Q_2)$ is finite, and $\Gamma(Q_1) = \infty$ if $\Gamma(Q_2) = \infty$. Continuity of the solution to the system of (A.25) and (A.26) with respect to the initial conditions implies that $\Gamma(Q)$ is continuous in Q at any point where $\Gamma(Q)$ is finite. Hence, $\Gamma(Q)$ is finite in an interval $[0, \bar{Q})$ and is infinite in $[\bar{Q}, \infty)$. Moreover, it is continuous an increasing in $Q \in [0, \bar{Q})$, and satisfies $\lim_{Q \to \bar{Q}} \Gamma(Q) = \infty$. Therefore, for large enough M, there exists a unique $\hat{Q} \in [0, \bar{Q})$ such that $\Gamma(\hat{Q}) = \gamma$. The solution to the system of (A.25) and (A.26) with initial conditions $R(\epsilon) = \max\{\gamma, K\}$ and $Q(\epsilon) = \hat{Q}$ has the properties in the lemma. ∎

Lemma A.4 (Bounds) *For any $\epsilon > 0$ and $M > \epsilon$ large enough, the solution to the system of (A.25) and (A.26) constructed in Lemma A.3 satisfies the following inequalities:*

- *If $\gamma < 1$, then $1 > R(w_t) > \gamma$ for all $w_t \in (\epsilon, M)$.*
- *If $\gamma > 1$, then $1 < R(w_t) < \gamma$ for all $w_t \in (\epsilon, M)$.*

Proof: Consider first the case $\gamma < 1$. To show that $R(w_t) < 1$ for all $w_t \in [\epsilon, M]$, we proceed by contradiction and assume that there exists $w_t \in (\epsilon, M)$ such that $R(w_t) \geq 1$. In that case, $R(w_t)$ reaches its maximum value over the closed interval $[\epsilon, M]$ at an interior point m, and $R(m) \geq 1$. Since m is interior, $R'(m) = 0$, and (A.26) implies

$$R(m)[1 - R(m)] + \frac{[\alpha m + R(m)]^2}{z} \left(\frac{Q(m) - r}{\rho - r} R(m) - 1 \right) = 0$$

$$\Rightarrow \frac{Q(m) - r}{\rho - r} R(m) - 1 = \frac{z R(m)[R(m) - 1]}{[\alpha m + R(m)]^2} \geq 0. \tag{A.36}$$

Differentiating (A.26) at m and using $R'(m) = 0$, we find

$$\begin{aligned} R''(m)m &= \frac{2\alpha[\alpha m + R(m)]}{z} \left(\frac{Q(m) - r}{\rho - r} R(m) - 1 \right) + \frac{[\alpha m + R(m)]^2}{z} \frac{Q'(m)R(m)}{\rho - r} \\ &= \frac{2\alpha[\alpha m + R(m)]}{z} \left(\frac{Q(m) - r}{\rho - r} R(m) - 1 \right) + \frac{[\alpha m + R(m)]^2}{z} \frac{Q(m)[R(m) - \gamma]R(m)}{(\rho - r)\gamma m}, \end{aligned} \tag{A.37}$$

where the second step follows from (A.25). The first term in (A.37) is non-negative because of (A.36). The second term is positive because $R(m) \geq 1 > \gamma$. Hence, $R''(m) > 0$, which contradicts m being an interior maximizer of $R(w_t)$.

To show that $R(w_t) > \gamma$ for all $w_t \in (\epsilon, M)$, we proceed by contradiction and assume that there exists $w_t \in (\epsilon, M)$ such that $R(w_t) \leq \gamma$. In that case, $R(w_t)$ takes its minimum value over the closed interval $[\epsilon, M]$ at an interior point m (and possibly at ϵ and M as well), and $R(m) \leq \gamma$. Suppose first that $R(m) = \gamma$. Proceeding as in (A.36) we find $\frac{Q(m)-r}{\rho-r}R(m) - 1 < 0$, and proceeding as in (A.37) we find $R''(m) < 0$, which contradicts m being an interior minimizer of $R(w_t)$. Suppose next that $R(m) < \gamma$. Since $R(\epsilon) \geq \gamma$ and $R(M) = \gamma$, there exist $(w_t, \hat{w}_t)$ such that $\epsilon \leq w_t < m < \hat{w}_t \leq M$ and $R(w_t) = R(\hat{w}_t) = \gamma$. The supremum m_1 and the infimum m_2 within the corresponding sets satisfy $m_1 < m < m_2$, $R(m_1) = R(m_2) = \gamma$, $R(w_t) < \gamma$ for all $w_t \in (m_1, m_2)$, $R'(m_1) \leq 0$ and $R'(m_2) \geq 0$. Using (A.26) and $R(m_1) = R(m_2) = \gamma$, we find that the inequalities $R'(m_1) \leq 0$ and $R'(m_2) \geq 0$ imply

$$\gamma(1-\gamma) + \frac{(\alpha m_1 + \gamma)^2}{z}\left(\frac{Q(m_1) - r}{\rho - r}\gamma - 1\right) \leq 0 \Rightarrow \frac{Q(m_1) - r}{\rho - r}\gamma - 1 \leq \frac{z\gamma(\gamma - 1)}{(\alpha m_1 + \gamma)^2}, \qquad \text{(A.38)}$$

$$\gamma(1-\gamma) + \frac{(\alpha m_2 + \gamma)^2}{z}\left(\frac{Q(m_2) - r}{\rho - r}\gamma - 1\right) \geq 0 \Rightarrow \frac{Q(m_2) - r}{\rho - r}\gamma - 1 \geq \frac{z\gamma(\gamma - 1)}{(\alpha m_2 + \gamma)^2}, \qquad \text{(A.39)}$$

respectively. Since $\gamma < 1$, (A.38) and (A.39) imply $Q(m_1) < Q(m_2)$. This is a contradiction because (A.26) and $R(w_t) < \gamma$ for all $w_t \in (m_1, m_2)$ imply $Q(m_1) > Q(m_2)$.

The inequalities in the case $\gamma > 1$ follow from similar arguments. Suppose, by contradiction, that there exists $w_t \in (\epsilon, M)$ such that $R(w_t) \geq \gamma$. In that case, $R(w_t)$ reaches its maximum value over the closed interval $[\epsilon, M]$ at an interior point m (and possibly at ϵ and M as well), and $R(m) \geq \gamma$. Proceeding as in (A.36) we find $\frac{Q(m)-r}{\rho-r}R(m) - 1 > 0$, and proceeding as in (A.37) we find $R''(m) > 0$, which contradicts m being an interior maximizer of $R(w_t)$.

Suppose next, by contradiction, that there exists $w_t \in (\epsilon, M)$ such that $R(w_t) \leq 1$. In that case, $R(w_t)$ reaches its minimum value over the closed interval $[\epsilon, M]$ at an interior point m, and $R(m) \leq 1$. Suppose first that $R(m) = 1$. Proceeding as in (A.36) we find $\frac{Q(m)-r}{\rho-r}R(m) - 1 = 0$, and proceeding as in (A.37) we find $R''(m) < 0$, which contradicts m being an interior minimizer of $R(w_t)$. Suppose next that $R(m) < 1$. Since $R(\epsilon) = R(M) = \gamma > 1$, there exist $(w_t, \hat{w}_t)$ such that $\epsilon < w_t < m < \hat{w}_t < M$ and $R(w_t) = R(\hat{w}_t) = 1$. The supremum m_1 and the infimum m_2 within the corresponding sets satisfy $m_1 < m < m_2$, $R(m_1) = R(m_2) = 1$, $R(w_t) < 1$ for all $w_t \in (m_1, m_2)$, $R'(m_1) \leq 0$ and $R'(m_2) \geq 0$. Using (A.26) and $R(m_1) = R(m_2) = 1$, we find that the inequalities

$R'(m_1) \leq 0$ and $R'(m_2) \geq 0$ imply

$$\frac{Q(m_1) - r}{\rho - r} - 1 \leq 0, \tag{A.40}$$

$$\frac{Q(m_2) - r}{\rho - r}\gamma - 1 \geq 0, \tag{A.41}$$

respectively, and hence $Q(m_1) \leq Q(m_2)$. This is a contradiction because (A.26) and $R(w_t) < 1 < \gamma$ for all $w_t \in (m_1, m_2)$ imply $Q(m_1) > Q(m_2)$. ∎

Lemma A.5 (Existence in $[0, \infty]$**)** *A solution to the system of (A.25) and (A.26), defined over the interval* $(0, \infty)$*, and such that* $\lim_{w_t \to 0} R(w_t) \in (0, \infty)$ *and* $\lim_{w_t \to \infty} Q(w_t) \in (0, \infty)$*, exists. This solution satisfies the inequalities in Lemma A.4 for all* $w_t \in (0, \infty)$.

Proof: We will construct the solution over $(0, \infty)$ as the simple limit of solutions over finite intervals $[\epsilon, M]$. We first derive the limit when M goes to infinity, holding $\epsilon > 0$ constant, and then derive the limit of those limits when ϵ goes to zero.

Fix $\epsilon > 0$ and denote by $(Q_{\epsilon,M}(w_t), R_{\epsilon,M}(w_t))$ the solution constructed in Lemma A.3 for $M > \epsilon$ large enough. Consider first the case $\gamma < 1$. Since for $M_2 > M_1$, $R_{\epsilon,M_2}(M_1) > \gamma$ (as implied by Lemma A.4) and $R_{\epsilon,M_1}(M_1) = \gamma$, Lemma A.2 implies $Q_{\epsilon,M_2}(\epsilon) > Q_{\epsilon,M_1}(\epsilon)$, and hence $Q_{\epsilon,M_2}(w_t) > Q_{\epsilon,M_1}(w_t)$ and $R_{\epsilon,M_2}(w_t) > R_{\epsilon,M_1}(w_t)$ for all $w_t \in (\epsilon, M_1]$. This means that the functions $M \to Q_{\epsilon,M}(w_t)$ and $M \to R_{\epsilon,M}(w_t)$, defined for given $w_t > \epsilon$ and for $M > w_t$, are increasing. The latter function is bounded above by one (as implied by Lemma A.4). The former function is also bounded above. Indeed, since $R_{\epsilon,M}(M) = \gamma$ and $R'_{\epsilon,M}(M) \leq 0$ (which follows from $R_{\epsilon,M}(M) = \gamma$ and $R_{\epsilon,M}(w_t) > \gamma$ for all $w_t \in (\epsilon, M)$), (A.26) implies

$$\frac{Q_{\epsilon,M}(M) - r}{\rho - r}\gamma - 1 \leq \frac{z\gamma(\gamma - 1)}{(\alpha M + \gamma)^2} < 0 \Rightarrow Q_{\epsilon,M}(M) < r + \frac{\rho - r}{\gamma}.$$

Since, in addition, $Q_{\epsilon,M}(w_t)$ is increasing in $w_t \in (\epsilon, M)$ (which follows from (A.25) and because $R_{\epsilon,M}(w_t) > \gamma$), $Q_{\epsilon,M}(w_t) < r + \frac{\rho-r}{\gamma}$ for all $w_t < M$. Being increasing and bounded above, the functions $M \to Q_{\epsilon,M}(w_t)$ and $M \to R_{\epsilon,M}(w_t)$ converge to limits when M goes to infinity. These limits, denoted by $Q_\epsilon(w_t)$ and $R_\epsilon(w_t)$, respectively, satisfy $r + \frac{\rho-r}{\gamma} \geq Q_\epsilon(w_t) \geq 0$ and $1 \geq R_\epsilon(w_t) \geq \gamma$ for all $w_t \in (\epsilon, \infty)$.

Consider next the case $\gamma < 1$. Since for $M_2 > M_1$, $R_{\epsilon,M_2}(M_1) < \gamma$ (as implied by Lemma A.4) and $R_{\epsilon,M_1}(M_1) = \gamma$, Lemma A.2 implies $Q_{\epsilon,M_2}(\epsilon) < Q_{\epsilon,M_1}(\epsilon)$, and hence $Q_{\epsilon,M_2}(w_t) < Q_{\epsilon,M_1}(w_t)$ and $R_{\epsilon,M_2}(w_t) < R_{\epsilon,M_1}(w_t)$ for all $w_t \in (\epsilon, M_1]$. This means that the functions $M \to Q_{\epsilon,M}(w_t)$ and $M \to R_{\epsilon,M}(w_t)$ are decreasing. The former function is bounded below by zero and the latter function is bounded below by one (as implied by Lemma A.4). Hence, the two functions converge to limits when M goes to infinity. These limits, denoted by $Q_\epsilon(w_t)$ and $R_\epsilon(w_t)$, respectively, satisfy $r + \frac{\rho-r}{\gamma}\left(1 + \frac{z(\gamma-1)}{\gamma}\right) \geq Q_\epsilon(w_t) \geq 0$ and $1 \leq R_\epsilon(w_t) \leq \gamma$ for all $w_t \in (\epsilon, \infty)$. The upper bound on $Q_\epsilon(w_t)$ can be derived by noting that since $R_{\epsilon,M}(\epsilon) = \gamma$ and $R'_{\epsilon,M}(\epsilon) \leq 0$ (which follows from $R_{\epsilon,M}(\epsilon) = \gamma$ and $R_{\epsilon,M}(w_t) < \gamma$ for all $w_t \in (\epsilon, M)$), (A.26) implies

$$\frac{Q_{\epsilon,M}(\epsilon) - r}{\rho - r}\gamma - 1 \leq \frac{z\gamma(\gamma-1)}{(\alpha\epsilon + \gamma)^2} \Rightarrow Q_{\epsilon,M}(\epsilon) < r + \frac{\rho - r}{\gamma}\left(1 + \frac{z(\gamma-1)}{\gamma}\right).$$

Since, in addition, $Q_{\epsilon,M}(w_t)$ is decreasing in $w_t \in (\epsilon, M)$ (which follows from (A.25) and because $R_{\epsilon,M}(w_t) < \gamma$), $Q_{\epsilon,M}(w_t) < r + \frac{\rho-r}{\gamma}\left(1 + \frac{z(\gamma-1)}{\gamma}\right)$ for all $w_t > \epsilon$.

We next show that the limits $Q_\epsilon(w_t)$ and $R_\epsilon(w_t)$, viewed as functions of w_t, are solutions to the system of (A.25) and (A.26). We denote by $Q^*_\epsilon(w_t)$ and $R^*_\epsilon(w_t)$ the limits of the functions $M \to Q'_{\epsilon,M}(w_t)$ and $M \to R'_{\epsilon,M}(w_t)$, respectively. These limits exist because the limits of $M \to Q_{\epsilon,M}(w_t)$ and $M \to R_{\epsilon,M}(w_t)$ exist, and because $(Q'_{\epsilon,M}(w_t), R'_{\epsilon,M}(w_t))$ are linked to $(Q_{\epsilon,M}(w_t), R_{\epsilon,M}(w_t))$ through (A.25) and (A.26). Consider $w_t \in (\epsilon, \infty)$ and a small neighborhood B around it. Since $(Q_{\epsilon,M}(m), R_{\epsilon,M}(m))$ are bounded uniformly for all $m \in B$ and for all M, (A.25) and (A.26) imply that the same is true for $(Q'_{\epsilon,M}(m), R'_{\epsilon,M}(m))$. The same is also true for $(Q''_{\epsilon,M}(m), R''_{\epsilon,M}(m))$, as can be seen by differentiating (A.25) and (A.26). Denoting the bound on $Q''_{\epsilon,M}(m)$ by $\bar{Q}_{\epsilon,B}$ and using the intermediate value theorem, we find that for all $m \in B$,

$$\left|\frac{Q_{\epsilon,M}(m) - Q_{\epsilon,M}(w_t)}{m - w_t} - Q'_{\epsilon,M}(w_t)\right| = \left|Q'_{\epsilon,M}(m') - Q'_{\epsilon,M}(w_t)\right| \leq \bar{Q}_{\epsilon,B}|m' - w_t| < \bar{Q}_{\epsilon,B}|m - w_t|, \tag{A.42}$$

where m' is between m and w_t. Taking limits in (A.42) when M goes to infinity, we find

$$\left|\frac{Q_\epsilon(m) - Q_\epsilon(w_t)}{m - w_t} - Q^*_\epsilon(w_t)\right| \leq \bar{Q}_{\epsilon,B}|m - w_t|. \tag{A.43}$$

Equation (A.43) implies that $Q_\epsilon(w_t)$ is differentiable in w_t and its derivative $Q'_\epsilon(w_t)$ is $Q^*_\epsilon(w_t)$. Likewise, $R_\epsilon(w_t)$ is differentiable in w_t and its derivative $R'_\epsilon(w_t)$ is $R^*_\epsilon(w_t)$. Taking limits in (A.25) and (A.26), written for $(Q_{\epsilon,M}(w_t), R_{\epsilon,M}(w_t))$, when M goes to infinity, we find that $(Q_\epsilon(w_t), R_\epsilon(w_t))$ are solutions to (A.25) and (A.26).

We next take the limits of $Q_\epsilon(w_t)$ and $R_\epsilon(w_t)$ when ϵ goes to zero, and show that these limits, denoted by $Q(w_t)$ and $R(w_t)$, respectively, are bounded and solve the system of (A.25) and (A.26). The steps parallel those when limits are taken for given ϵ and for M going to infinity, but additional complications arise. Consider first the case $K < \gamma < 1$. Since for $\epsilon_2 < \epsilon_1$ and for M large enough, $R_{\epsilon_2,M}(\epsilon_1) > \gamma$ (as implied by Lemma A.4) and $R_{\epsilon_1,M}(\epsilon_1) = \gamma$, Lemma A.2 (applied with M rather than ϵ as starting value) implies $Q_{\epsilon_2,M}(M) < Q_{\epsilon_1,M}(M)$, and hence $Q_{\epsilon_2,M}(w_t) < Q_{\epsilon_1,M}(w_t)$ and $R_{\epsilon_2,M}(w_t) > R_{\epsilon_1,M}(w_t)$ for all $w_t \in [\epsilon_1, M)$. Taking limits when M goes to infinity, we find $Q_{\epsilon_2}(w_t) \leq Q_{\epsilon_1}(w_t)$ and $R_{\epsilon_2}(w_t) \geq R_{\epsilon_1}(w_t)$ for all $w_t \in (\epsilon_1, \infty)$. Hence, the functions $\epsilon \to Q_\epsilon(w_t)$ and $\epsilon \to R_\epsilon(w_t)$, defined for given $w_t > \epsilon$, are increasing and decreasing, respectively. Their limits and bounds are constructed as in the first part of the proof (M goes to infinity).

Consider next the case $\gamma > 1$. Since for $\epsilon_2 < \epsilon_1$ and for M large enough, $R_{\epsilon_2,M}(\epsilon_1) < \gamma$ (as implied by Lemma A.4) and $R_{\epsilon_1,M}(\epsilon_1) = \gamma$, Lemma A.2 implies $Q_{\epsilon_2,M}(M) > Q_{\epsilon_1,M}(M)$, and hence $Q_{\epsilon_2,M}(w_t) > Q_{\epsilon_1,M}(w_t)$ and $R_{\epsilon_2,M}(w_t) < R_{\epsilon_1,M}(w_t)$ for all $w_t \in [\epsilon_1, M)$. Taking limits when M goes to infinity, we find $Q_{\epsilon_2}(w_t) \geq Q_{\epsilon_1}(w_t)$ and $R_{\epsilon_2}(w_t) \leq R_{\epsilon_1}(w_t)$ for all $w_t \in (\epsilon_1, \infty)$. Hence, the functions $\epsilon \to Q_\epsilon(w_t)$ and $\epsilon \to R_\epsilon(w_t)$, defined for given $w_t > \epsilon$, are decreasing and increasing, respectively. Their limits and bounds are constructed as in the first part of the proof.

Consider finally the case $\gamma < K$. This case is trickier: unlike when $K < \gamma < 1$, $R_{\epsilon_2,M}(\epsilon_1)$ for $\epsilon_2 < \epsilon_1$ can be larger or smaller than $\max\{\gamma, K\}$, and hence can be larger or smaller than $R_{\epsilon_1,M}(\epsilon_1)$. The key observation is that if there exist $\epsilon_2 < \epsilon_1$ such that $R_{\epsilon_2,M}(\epsilon_1)$ is larger than $\max\{\gamma, K\}$, then the function $\epsilon \to R_{\epsilon,M}(\epsilon_1)$ is decreasing for $\epsilon < \epsilon_2$. Indeed, if that function were not decreasing, then there would exist $\epsilon_4 < \epsilon_3 \leq \epsilon_2$ such that $R_{\epsilon_3,M}(w_t)$ and $R_{\epsilon_4,M}(w_t)$ would cross twice, for a $w_t < M$ and for $w_t = M$, contradicting Lemma A.2. Hence, there are two possibilities for ϵ close to zero: either the function $\epsilon \to R_{\epsilon,M}(\epsilon_1)$ is decreasing, or it is increasing. In either case, we follow the first part of the proof to construct the limits and bounds. The argument that the limits $Q(w_t)$ and $R(w_t)$, viewed as functions of w_t, are solutions to the system of (A.25) and (A.26) is as in the first part of the proof, for all values of γ.

We next show that $\lim_{w_t \to 0} R(w_t) \in (0, \infty)$ and $\lim_{w_t \to \infty} Q(w_t) \in (0, \infty)$. Consider first the limits at zero. When $\gamma < 1$, $R(w_t) \geq \gamma$ and hence (A.25) implies that $Q(w_t)$ is increasing. When $\gamma < 1$, $R(w_t) \leq \gamma$ and hence (A.25) implies that $Q(w_t)$ is decreasing. In both cases, $Q(w_t)$ is monotone, and because it is bounded it converges to a finite limit when w_t goes to zero. Suppose, by contradiction, that $\lim_{w_t \to 0} R(w_t)$ does not exist. This means that as w_t goes to zero $R(w_t)$ oscillates more and more rapidly between maxima and minima whose distance is bounded away from zero. Since $R'(w_t) = 0$ at the maxima, (A.26) implies

$$1 - \limsup_{w_t \to 0} R(w_t) - \frac{\limsup_{w_t \to 0} R(w_t)}{z} \left(\frac{\lim_{w_t \to 0} Q(w_t) - r}{\rho - r} \limsup_{w_t \to 0} R(w_t) + 1 \right) = 0, \quad \text{(A.44)}$$

after taking the limit along a sequence of maxima as w_t goes to zero. Likewise, since $R'(w_t) = 0$ at the minima, (A.26) implies

$$1 - \liminf_{w_t \to 0} R(w_t) - \frac{\liminf_{w_t \to 0} R(w_t)}{z} \left(\frac{\lim_{w_t \to 0} Q(w_t) - r}{\rho - r} \liminf_{w_t \to 0} R(w_t) + 1 \right) = 0, \quad \text{(A.45)}$$

after taking the limit along a sequence of minima as w_t goes to zero. Equations (A.44) and (A.45) imply $\limsup_{w_t \to 0} R(w_t) = \liminf_{w_t \to 0} R(w_t)$, a contradiction. Since $R(w_t) \geq \min\{\gamma, 1\}$, $\lim_{w_t \to 0} R(w_t) \in (0, \infty)$. Consider next the limits at infinity. Since $Q(w_t)$ is bounded and monotone, it converges to a finite limit, which is non-negative because $Q(w_t)$ is non-negative. If $\lim_{w_t \to \infty} Q(w_t) = 0$, then (A.26) and $\max\{\gamma, 1\} \geq R(w_t) \geq \min\{\gamma, 1\}$ for all $w_t \in (0, \infty)$ imply $\lim_{w_t \to \infty} R'(w_t) = -\infty$, which contradicts $R(w_t) \geq \min\{\gamma, 1\}$. Hence, $\lim_{w_t \to \infty} Q(w_t) \in (0, \infty)$.

We finally show that the inequalities in Lemma A.4 hold. The steps are similar to those in Lemma A.4, after replacing ϵ by zero and M by infinity, and noting that the boundary conditions for $R(w_t)$ remain the same because of Lemma A.1. The only change concerns the proof that $R(w_t) > \gamma$ for $\gamma < 1$. If in that proof the supremum m_1 is zero, $\lim_{w_t \to 0} R'(w_t) w_t = 0$ implies that (A.38) holds as an equality. ∎

Lemma A.6 (Decreasing $A(w_t)$) *For the solution to the system of (A.25) and (A.26) constructed in Lemma A.5, $A(w_t)$ is decreasing. The same property holds for the solution constructed in Lemma A.3, if $\gamma < 1$. If $\gamma > 1$, then the same property holds if, in addition, M is large enough and $w_t \in [\epsilon, M_0]$ for any $M_0 > \epsilon$ which is kept fixed as M grows.*

Proof: We first show the property for the solution to the system of (A.25) and (A.26) constructed in Lemma A.5. We next extend the proof to the solution constructed in Lemma A.3.

Suppose, by contradiction, that there exist $m_2 > m_1 > 0$ such that $A(m_1) \leq A(m_2)$. If $A(m_1) < A(m_2)$, then there exists $m_3 \in (m_1, m_2)$ such that $A'(m_3) > 0$. Consider the set of $w_t < m_3$ such that $A'(\hat{w}_t) > 0$ for all $\hat{w}_t \in (w_t, m_3)$. The infimum m within that set is strictly smaller than m_3 since $A'(w_t) > 0$ for all w_t close to m_3. It is also positive because $\lim_{w_t \to 0} A(w_t) = \infty$ (which follows from $\lim_{w_t \to 0} R(w_t) \in (0, \infty)$), and it satisfies $A'(m) = 0$ and $A''(m) \geq 0$. Consider next the set of $w_t > m_3$ such that $A'(\hat{w}_t) > 0$ for all $\hat{w}_t \in (m_3, w_t)$. The supremum $\hat{m}$ within that set is strictly larger than m_3. It is also finite because $\lim_{w_t \to 0} A(w_t) = 0$ (which follows from $\lim_{w_t \to 0} R(w_t) \in (0, \infty)$), and it satisfies $A'(\hat{m}) = 0$ and $A''(\hat{m}) \leq 0$. Moreover, by the definition of m and $\hat{m}$, $A'(w_t) > 0$ for all $w_t \in (m, \hat{m})$.

If $A(m_1) = A(m_2)$, then points $\hat{m} > m > 0$ with the properties $A'(m) = A'(\hat{m}) = 0$, $A''(m) \geq 0$, $A''(\hat{m}) \leq 0$, and $A'(w_t) > 0$ for all $w_t \in (m, \hat{m})$, can again be constructed, except that the latter inequality can be weak. To perform the construction, we distinguish the case where there exists $w_t \in (m_1, m_2)$ such that $A(w_t) \neq A(m_1) = A(m_2)$ and replace m_1 or m_2 by w_t, and the case where $A(w_t) = A(m_1) = A(m_2)$ for all $w_t \in (m_1, m_2)$ and take $m = m_1$ and $\hat{m} = m_2$.

Using $Q(w_t)$ instead of $q(w_t)$, we can write (A.24) as

$$A'(w_t) = -A(w_t)^2 + \frac{[\alpha + A(w_t)]^2}{z} \left(\frac{Q(w_t) - r}{\rho - r} A(w_t) w_t - 1 \right). \tag{A.46}$$

Differentiating (A.46) at a point where $A'(w_t) = 0$, and using (A.25), we find

$$A''(w_t) = \frac{[\alpha + A(w_t)]^2}{z} \left(\frac{Q(w_t) \frac{R(w_t) - \gamma}{\gamma w_t}}{\rho - r} A(w_t) w_t + \frac{Q(w_t) - r}{\rho - r} A(w_t) \right).$$

Hence, the sign of $A''(w_t)$ at such a point is the same as of

$$H(w_t) \equiv Q(w_t) \frac{R(w_t) - \gamma}{\gamma} + Q(w_t) - r = \frac{Q(w_t) R(w_t)}{\gamma} - r. \tag{A.47}$$

Using (A.25) and (A.26), we find that for $w_t \in (m, \hat{m})$,

$$\begin{aligned}
H'(w_t) &= \frac{Q'(w_t)R(w_t) + Q(w_t)R'(w_t)}{\gamma} \\
&= \frac{Q(w_t)}{\gamma w_t}\left[\frac{[R(w_t)-\gamma]R(w_t)}{\gamma} + R(w_t)[1-R(w_t)] + \frac{[\alpha w_t + R(w_t)]^2}{z}\left(\frac{Q(w_t)-r}{\rho - r}R(w_t) - 1\right)\right] \\
&\geq \frac{Q(w_t)}{\gamma w_t}\left[\frac{[R(w_t)-\gamma]R(w_t)}{\gamma} + R(w_t)[1-R(w_t)] + R(w_t)^2\right] \\
&= \frac{Q(w_t)R(w_t)^2}{\gamma^2 w_t} > 0,
\end{aligned}$$

where the third step follows by using $A'(w_t) \geq 0$ for $w_t \in (m, \hat{m})$ and (A.46). Since $A''(m) \geq 0$, $H(m) \geq 0$. Since $H(w_t)$ is increasing in $(m, \hat{m})$, $H(\hat{m}) > 0$, and hence $A''(\hat{m}) > 0$, a contradiction.

We next extend the proof to the solution constructed in Lemma A.3. Suppose that $\gamma < 1$. Since $R(\epsilon) = \max\{\gamma, K\} \geq R(M) = \gamma$, there exists $m_\epsilon > \epsilon$ such that $R(m_\epsilon) = R(\epsilon)$ and $R'(m_\epsilon) \leq 0$. Since $R'(m_\epsilon) \leq 0$, (A.26) implies

$$\frac{Q(m_\epsilon) - r}{\rho - r}R(m_\epsilon) - 1 < 0 \Rightarrow \frac{Q(\epsilon) - r}{\rho - r}R(m_\epsilon) - 1 < 0,$$

where the second step follows because $Q(w_t)$ is increasing for $\gamma < 1$. Equation (A.46) then implies $A'(\epsilon) < 0$. Since, $R(w_t) > R(M) = \gamma$ for all $w_t \in (\epsilon, M)$, $R'(M) \leq 0$ and hence $A'(M) < 0$. Using $A'(\epsilon) < 0$ and $A'(M) < 0$, we can define m and $\hat{m}$ by proceeding as above, and derive a contradiction. Suppose next that $\gamma > 1$. Since $R(w_t) < R(\epsilon) = \gamma$ for all $w_t \in (\epsilon, M)$, $R'(\epsilon) \leq 0$ and hence $A'(\epsilon) < 0$. Since, in addition, for any given w_t, the values of $Q(w_t)$ and $R(w_t)$ under the solution over $[\epsilon, M]$ converge to the values under the solution over $[\epsilon, \infty)$ when M goes to infinity, $A(M_0)$ is close to its positive limit under the solution over $[\epsilon, \infty)$ for M large enough. It is, therefore, larger than $A(M) = \frac{\gamma}{M}$ for M large enough. Using $A'(\epsilon) < 0$ and $A(M_0) > A(M)$, we can define m and $\hat{m}$ by proceeding as above, and derive a contradiction. ∎

The results in Theorem 3.1 for the case of short-lived hedgers follow from the lemmas proved so far. Lemma A.5 yields the existence of a solution with the required boundary conditions. The same lemma yields the comparisons between $A(w_t)$, $\frac{1}{w_t}$ and $\frac{\gamma}{w_t}$. The monotonicity of $q(w_t)$ follows from these comparisons and (A.23). Lemma A.6 yields the monotonicity of $A(w_t)$. Lemma A.1 yields the limits at zero and infinity.

We next turn to the case where hedgers are long-lived. The system of ODEs consists of (A.25),

$$R'(w_t)w_t = R(w_t)\left[1 - R(w_t)\right] + \frac{\left[\alpha w_t + R(w_t) - F'(w_t)w_t\right]^2}{z}\left(\frac{Q(w_t) - r}{\rho - r}R(w_t) - 1\right), \quad \text{(A.48)}$$

which replaces (A.26), and

$$\begin{aligned}1 =& \frac{rF(w_t) - r\log(r) - \alpha u^\top \bar{D} + \rho - \bar{\rho}}{\rho - r} + \frac{Q(w_t) - r}{\rho - r}F'(w_t)w_t \\ &- \frac{z\left\{F''(w_t)w_t^2 - R(w_t)\left[2\alpha w_t + R(w_t) - 2F'(w_t)w_t\right]\right\}}{\left[\alpha w_t + R(w_t) - F'(w_t)w_t\right]^2},\end{aligned} \quad \text{(A.49)}$$

which is (3.22) written in terms of $Q(w_t) = q(w_t)^{-\frac{1}{\gamma}}$ and $R(w_t) = A(w_t)w_t$.

The analysis in Lemma A.1 on the limits of $Q(w_t)$ and $R(w_t)$ at zero and infinity carries through. For the limits at zero, we need to observe additionally that because $\lim_{w_t\to 0} F(w_t)$ is finite, the limit of $F'(w_t)w_t$ (which is assumed to exist) is zero. For the limits at infinity, we need to observe additionally that because $\lim_{w_t\to\infty} F(w_t)$ is finite, the limit of $F'(w_t)w_t$ (which is assumed to exist) is zero. Lemma A.1 provides the argument for $\lim_{w_t\to 0} F'(w_t)w_t$ (in the case of $R(w_t)$). The argument for $\lim_{w_t\to\infty} F'(w_t)w_t$ is similar: if $\lim_{w_t\to\infty} F'(w_t)w_t \neq 0$, then $|F'(w_t)| \geq \frac{\ell}{w_t}$ for $w_t > M$ and for positive M and ℓ. Since, however, for $w_t > M$

$$F(w_t) = F(M) + \int_M^{w_t} F'(\hat{w}_t)d\hat{w}_t \Rightarrow |F(w_t) - F(M)| \geq \int_M^{w_t} \frac{\ell}{\hat{w}_t}d\hat{w}_t = \ell\log\left(\frac{w_t}{M}\right),$$

$\lim_{w_t\to\infty} R(w_t)$ would be plus or minus infinity, a contradiction. Since the limits of $F'(w_t)w_t$ at zero and infinity are zero, the limits of $[A(w_t) - F'(w_t)]w_t$ at zero and infinity are the same as those of $A(w_t)w_t$, and hence are the same as in Lemma A.1.

The limits of $F(w_t)$ at zero and infinity follow from (A.49), the limits of $Q(w_t)$ and $R(w_t)$, $\lim_{w_t\to 0} F'(w_t)w_t = 0$, $\lim_{w_t\to\infty} F'(w_t)w_t = 0$, $\lim_{w_t\to 0} F''(w_t)w_t^2 = 0$ and $\lim_{w_t\to\infty} F''(w_t)w_t^2 = 0$. The latter two properties follow using similar arguments as for $F'(w_t)w_t$. If $\lim_{w_t\to 0} F''(w_t)w_t^2 \neq 0$, then $|F''(w_t)| \geq \frac{\ell}{w_t^2}$ for $w_t < \epsilon$ and for positive ϵ and ℓ. Since, however, for $w_t < \epsilon$

$$\begin{aligned}&F(w_t) = F(\epsilon) + F'(\epsilon)(w_t - \epsilon) + \int_\epsilon^{w_t} F''(\hat{w}_t)(w_t - \hat{w}_t)d\hat{w}_t \\ &\Rightarrow |F(w_t) - F(\epsilon) - F'(\epsilon)(w_t - \epsilon)| \geq \int_{w_t}^{\epsilon} \frac{\ell}{\hat{w}_t^2}(\hat{w}_t - w_t)d\hat{w}_t = \ell\left[\log\left(\frac{\epsilon}{w_t}\right) - \frac{\epsilon - w_t}{\epsilon}\right],\end{aligned}$$

$\lim_{w_t \to 0} F(w_t)$ would be plus or minus infinity, a contradiction. The argument for $\lim_{w_t \to \infty} F''(w_t) w_t^2 = 0$ is a similar adaptation of that used to establish $\lim_{w_t \to \infty} F'(w_t) w_t = 0$. ∎

Proof of Corollary 3.1: Part (i) follows from $A(w_t)$ being decreasing in w_t, $F(w_t) = 0$ and (3.18). Part (ii) follows from the former two properties, (3.1) and (3.17). Part (iii) follows from the same two properties and because (3.17) implies that the market prices of risk are given by

$$\eta_t \equiv \left(\sigma^\top\right)^{-1} (\bar{D} - \pi_t) = \frac{\alpha A(w_t)}{\alpha + A(w_t) - F'(w_t)} \sigma u. \tag{A.50}$$

Part (iv) follows from the same two properties and (3.19). ∎

We next prove two useful lemmas on the monotonicity of $R(w_t)$ and $\alpha w_t + R(w_t)$.

Lemma A.7 (Monotonicity of $R(w_t)$) *For the solution to the system of (A.25) and (A.26) constructed in Lemma A.5, and for that constructed in Lemma A.3:*

- *If $\gamma < K$, then $R(w_t)$ is either decreasing for all values of w_t, or is hump-shaped.*
- *If $K < \gamma < 1$, then $R(w_t)$ is hump-shaped.*
- *If $\gamma > 1$, then $R(w_t)$ is inverse hump-shaped.*

Proof: We show the monotonicity properties for the solution to the system of (A.25) and (A.26) constructed in Lemma A.5. The proof for the solution constructed in Lemma A.3 follows by the same arguments and because the limits of $R(w_t)$ at zero and infinity under the former solution are the same as the boundary conditions at ϵ and M, respectively, under the latter solution.

Consider first the case $K < \gamma < 1$. Since $\lim_{w_t \to 0} R(w_t) = \lim_{w_t \to \infty} R(w_t) = \gamma$ and $1 > R(w_t) > \gamma$ for all $w_t \in (0, \infty)$, there exists $w_t \in (0, \infty)$ at which $R(w_t)$ is maximized. If $R(w_t)$ is not hump-shaped, then there exist $m_1 < m_2 < m_3$ such that $R(m_1) > R(m_2)$ and $R(m_2) < R(m_3)$. We can then choose $\xi \in (\gamma, 1)$ and $(\hat{m}_i, \hat{\hat{m}}_i)$ for $i = 1, 2, 3$, such that $\hat{m}_1 < m_1 < \hat{\hat{m}}_1 \leq \hat{m}_2 < m_2 < \hat{\hat{m}}_2 \leq \hat{m}_3 < m_3 < \hat{\hat{m}}_3$, $R(\hat{m}_i) = R(\hat{\hat{m}}_i) = \xi$ for $i = 1, 2, 3$, $R'(\hat{m}_i) > 0$ and $R'(\hat{\hat{m}}_i) < 0$ for $i = 1, 3$, $R'(\hat{m}_2) < 0$ and $R'(\hat{\hat{m}}_2) > 0$, $R(w_t) > \xi$ for all $w_t \in (\hat{m}_i, \hat{\hat{m}}_i)$ for $i = 1, 3$, and $R(w_t) < \xi$

for all $w_t \in (\hat{m}_2, \hat{\hat{m}}_2)$. Equation (A.26) implies that $\hat{R}_\xi(\hat{m}_i) > 0$ and $\hat{R}_\xi(\hat{\hat{m}}_i) < 0$ for $i = 1, 3$, and $\hat{R}_\xi(\hat{m}_2) < 0$ and $\hat{R}_\xi(\hat{\hat{m}}_2) > 0$, where the function $\hat{R}_\xi(w_t)$ is defined by

$$\hat{R}_\xi(w_t) \equiv \xi(1-\xi) + \frac{(\alpha w_t + \xi)^2}{z}\left(\frac{Q(w_t) - r}{\rho - r}\xi - 1\right). \tag{A.51}$$

Therefore, there exist $m_1' \in (\hat{m}_1, \hat{\hat{m}}_1)$, $m_2' \in (\hat{m}_2, \hat{\hat{m}}_2)$ and $m_3' \in (\hat{m}_3, \hat{\hat{m}}_3)$ such that $\hat{R}_\xi(m_i') = 0$ for $i = 1, 2, 3$, $\hat{R}_\xi'(m_1') \le 0$, $\hat{R}_\xi'(m_2') \ge 0$ and $\hat{R}_\xi'(m_3') \le 0$. Differentiating (A.51) and using (A.26) and $R(\hat{m}_i) = R(\hat{\hat{m}}_i) = \xi$ for $i = 1, 2, 3$, we find

$$\hat{R}_\xi'(w_t) = \frac{2\alpha(\alpha w_t + \xi)}{z}\left(\frac{Q(w_t) - r}{\rho - r}\xi - 1\right) + \frac{(\alpha w_t + \xi)^2}{z}\frac{Q(w_t)[R(w_t) - \gamma]}{(\rho - r)\gamma w_t}\xi. \tag{A.52}$$

If $\hat{R}_\xi(w_t) = 0$, then we can write (A.52) as

$$\hat{R}_\xi'(w_t) = -\frac{2\alpha\xi(1-\xi)}{\alpha w_t + \xi} + \frac{(\alpha w_t + \xi)^2}{z}\frac{\left[r + (\rho - r)\left(\frac{1}{\xi} - \frac{z(1-\xi)}{(\alpha w_t + \xi)^2}\right)\right][R(w_t) - \gamma]}{(\rho - r)\gamma w_t}\xi. \tag{A.53}$$

Since $\hat{R}_\xi(m_i') = 0$ and $\hat{R}_\xi'(m_i') \le 0$ for $i = 1, 3$, (A.53) implies

$$\frac{(\alpha m_i' + \xi)^3}{2\alpha z}\frac{\left[r + (\rho - r)\left(\frac{1}{\xi} - \frac{z(1-\xi)}{(\alpha m_i' + \xi)^2}\right)\right][R(m_i') - \gamma]}{(1-\xi)(\rho - r)\gamma m_i'} \le 1 \tag{A.54}$$

for $i = 1, 3$. Likewise, since $\hat{R}_\xi(m_2') = 0$ and $\hat{R}_\xi'(m_2') \ge 0$, (A.53) implies

$$\frac{(\alpha m_2' + \xi)^3}{2\alpha z}\frac{\left[r + (\rho - r)\left(\frac{1}{\xi} - \frac{z(1-\xi)}{(\alpha m_2' + \xi)^2}\right)\right][R(m_2') - \gamma]}{(1-\xi)(\rho - r)\gamma m_2'} \ge 1. \tag{A.55}$$

Consider the function

$$\begin{aligned} H_\xi(w_t) \equiv& \frac{(\alpha w_t + \xi)^3}{w_t}\left[r + (\rho - r)\left(\frac{1}{\xi} - \frac{z(1-\xi)}{(\alpha w_t + \xi)^2}\right)\right] \\ =& \left(r + \frac{\rho - r}{\xi}\right)(\alpha^3 w_t^2 + 3\alpha^2\xi w_t + 3\alpha\xi^2) - (\rho - r)z(1-\xi)\alpha + h_\xi\frac{\xi}{w_t}, \end{aligned}$$

where

$$h_\xi \equiv \left[\left(r + \frac{\rho - r}{\xi}\right)\xi^2 - (\rho - r)z(1-\xi)\right].$$

If $h_\xi \leq 0$, then $H_\xi(w_t)$ is increasing in w_t. Since $\gamma < R(m_2') < \xi$ (because $m_2' \in (\hat{m}_2, \hat{\hat{m}}_2)$) and $R(m_3') > \xi > \gamma$ (because $m_3' \in (\hat{m}_3, \hat{\hat{m}}_3)$), (A.55) implies that (A.54) for $i = 3$ should hold as a strict inequality in the opposite direction, a contradiction. If $h_\xi > 0$, then $H_\xi(w_t)$ is inverse hump-shaped in w_t. If $H_\xi(m_3') \geq H_\xi(m_2')$, then the previous argument yields a contradiction. Therefore, $H_\xi(m_3') < H_\xi(m_2')$, which means that m_2' is in the decreasing part of $H_\xi(w_t)$ and hence $H_\xi(m_1') > H_\xi(m_2')$. Since $R(m_1') > \xi > \gamma$ (because $m_1' \in (\hat{m}_1, \hat{\hat{m}}1)$), (A.54) for $i = 1$ implies that (A.55) should hold as a strict inequality in the opposite direction, a contradiction.

Consider next the case $\gamma > 1$. Since $\lim_{w_t \to 0} R(w_t) = \lim_{w_t \to \infty} R(w_t) = \gamma$ and $1 < R(w_t) < \gamma$ for all $w_t \in (0, \infty)$, there exists $w_t \in (0, \infty)$ at which $R(w_t)$ is minimized. If $R(w_t)$ is not inverse hump-shaped, then there exist $m_1 < m_2 < m_3$ such that $R(m_1) < R(m_2)$ and $R(m_2) > R(m_3)$. We can then choose $\xi \in (1, \gamma)$ and $(\hat{m}_i, \hat{\hat{m}}_i)$ for $i = 1, 2, 3$, such that $\hat{m}_1 < m_1 < \hat{\hat{m}}_1 \leq \hat{m}_2 < m_2 < \hat{\hat{m}}_2 \leq \hat{m}_3 < m_3 < \hat{\hat{m}}_3$, $R'(\hat{m}_i) < 0$ and $R'(\hat{\hat{m}}_i) > 0$ for $i = 1, 3$, $R'(\hat{m}_2) > 0$ and $R'(\hat{\hat{m}}_2) < 0$, $R(w_t) < \xi$ for all $w_t \in (\hat{m}_i, \hat{\hat{m}}_i)$ for $i = 1, 3$, and $R(w_t) > \xi$ for all $w_t \in (\hat{m}_2, \hat{\hat{m}}_2)$. Equation (A.26) implies that $\hat{R}_\xi(\hat{m}_i) < 0$ and $\hat{R}_\xi(\hat{\hat{m}}_i) > 0$ for $i = 1, 3$, and $\hat{R}_\xi(\hat{m}_2) > 0$ and $\hat{R}_\xi(\hat{\hat{m}}_2) < 0$. Therefore, there exist $m_1' \in (\hat{m}_1, \hat{\hat{m}}_1)$, $m_2' \in (\hat{m}_2, \hat{\hat{m}}_2)$ and $m_3' \in (\hat{m}_3, \hat{\hat{m}}_3)$ such that $\hat{R}_\xi(m_i') = 0$ for $i = 1, 2, 3$, $\hat{R}_\xi'(m_1') \leq 0$, $\hat{R}_\xi'(m_2') \geq 0$ and $\hat{R}_\xi'(m_3') \leq 0$. Equations (A.54) and (A.55) still hold, with both $R(m_i) - \gamma$ and $1 - \xi$ being negative. If $h_\xi \leq 0$, in which case $H_\xi(w_t)$ is increasing in w_t, then $\gamma > R(m_2') > \xi$, $R(m_3') < \xi < \gamma$, and (A.55) imply that (A.54) for $i = 3$ should hold as a strict inequality in the opposite direction, a contradiction. If $h_\xi > 0$, in which case $H_\xi(w_t)$ is inverse hump-shaped in w_t, then the previous argument implies that $H_\xi(m_3') < H_\xi(m_2')$, which means that m_2' is in the decreasing part of $H_\xi(w_t)$ and hence $H_\xi(m_1') > H_\xi(m_2')$. Since $R(m_1') < \xi < \gamma$, (A.54) for $i = 1$ implies that (A.55) should hold as a strict inequality in the opposite direction, a contradiction.

Consider finally the case $\gamma < K$. Since $\lim_{w_t \to 0} R(w_t) = K$ and $\lim_{w_t \to \infty} R(w_t) = \gamma$, $R(w_t)$ is maximized at its zero limit or at a $w_t \in (0, \infty)$. If $R(w_t)$ is not decreasing or hump-shaped, then there exist $m_1 < m_2 < m_3$ such that $R(m_1) > R(m_2)$ and $R(m_2) < R(m_3)$. If $R(m_2) \geq K$, then we can proceed as in the case $K < \gamma < 1$ to find a contradiction. If $R(m_2) < K$, then we can choose $\xi \in (\gamma, K)$ and $(\hat{m}_i, \hat{\hat{m}}_i)$ for $i = 2, 3$, such that $\hat{m}_2 < m_2 < \hat{\hat{m}}_2 \leq \hat{m}_3 < m_3 < \hat{\hat{m}}_3$, $R'(\hat{m}_2) < 0$ and $R'(\hat{\hat{m}}_2) > 0$, $R'(\hat{m}_3) > 0$ and $R'(\hat{\hat{m}}_3) < 0$, $R(w_t) < \xi$ for all $w_t \in (\hat{m}_2, \hat{\hat{m}}_2)$, and $R(w_t) > \xi$ for all $w_t \in (\hat{m}_3, \hat{\hat{m}}_3)$. (Note that $(\hat{m}_1, \hat{\hat{m}}_1)$ may not exist because $\xi < K = \lim_{w_t \to 0} R(w_t)$.) Equation

(A.26) implies that $\hat{R}_\xi(\hat{m}_2) < 0$, $\hat{R}_\xi(\hat{\hat{m}}_2) > 0$, $\hat{R}_\xi(\hat{m}_3) > 0$ and $\hat{R}_\xi(\hat{\hat{m}}_3) < 0$. Therefore, there exist $m_2' \in (\hat{m}_2, \hat{\hat{m}}_2)$ and $m_3' \in (\hat{m}_3, \hat{\hat{m}}_3)$ such that $\hat{R}_\xi(m_i') = 0$ for $i = 2, 3$, $\hat{R}_\xi'(m_2') \geq 0$ and $\hat{R}_\xi'(m_3') \leq 0$. Equations (A.54) for $i = 3$ and (A.55) still hold. Moreover, $h_\xi < 0$ because h_ξ is increasing in ξ, $\xi < K$, and $h_K = 0$ as implied by (3.23). Therefore, $H_\xi(w_t)$ is increasing in w_t, and (A.55) implies that (A.54) for $i = 3$ should hold as a strict inequality in the opposite direction, a contradiction. ∎

Lemma A.8 ($\alpha w_t + R(w_t)$ increasing) *For the solution to the system of (A.25) and (A.26) constructed in Lemma A.5, $\alpha w_t + R(w_t)$ is increasing if $\gamma \leq 1$. The same property holds for the solution constructed in Lemma A.3, provided that ϵ and M are small and large enough, respectively, and that $w_t \in [\epsilon, M_0]$ for any $M_0 > \epsilon$ which is kept fixed as ϵ shrinks and M grows.*

Proof: We first show the property for the solution to the system of (A.25) and (A.26) constructed in Lemma A.5. We next extend the proof to the solution constructed in Lemma A.3. We assume $\gamma < 1$: for $\gamma = 1$, the property trivially holds because $R(w_t) = 1$.

Consider first the case $K < \gamma < 1$. Since $R(w_t)$ is hump-shaped (Lemma A.7), $R'(w_t) > -\alpha$ for all $w_t \in (0, \hat{w}_t)$, where $\hat{w}_t$ exceeds the maximizer of $R(w_t)$. Denote by m the supremum of the set of $\hat{w}_t$ such that $R'(w_t) > -\alpha$ for all $w_t \in (0, \hat{w}_t)$. If m is infinite, then $\alpha w_t + R(w_t)$ is increasing. Suppose, by contradiction, that m is finite. By its definition, m satisfies $R'(m) = -\alpha$, and since $R'(w_t) > -\alpha$ for all $w_t < m$, it also satisfies $R''(m) \leq 0$. Differentiating (A.26) at m and using $R'(m) = -\alpha$, we find

$$R''(m)m - \alpha = -\alpha[1 - 2R(m)] + \frac{[\alpha m + R(m)]^2}{z} \left. \frac{d\left(\frac{Q(w_t)-r}{\rho-r} R(w_t)\right)}{dw_t} \right|_{w_t=m}$$

$$\Rightarrow R''(m)m = 2\alpha R(m) + \frac{[\alpha m + R(m)]^2}{z(\rho - r)} L'(m), \tag{A.56}$$

where $L(w_t) \equiv [Q(w_t) - r]R(w_t)$. To derive a contradiction, it suffices to show that $L'(m) \geq 0$, since (A.56) would then imply $R''(m) > 0$.

Using (A.25), we find

$$L'(w_t) = Q(w_t) \frac{R(w_t) - \gamma}{\gamma w_t} R(w_t) + [Q(w_t) - r]R'(w_t). \tag{A.57}$$

Equations (A.57), $R(w_t) > \gamma$ for all $w_t \in (0, \infty)$, and $R'(m) = -\alpha < 0$ imply $L'(m) \geq 0$ if $Q(m) \leq r$. To show that $L'(m) \geq 0$ also if $Q(m) > r$, we proceed by contradiction, and suppose that $L'(m) < 0$. Since $R'(m) = -\alpha < 0$, (A.26) implies

$$\frac{Q(m) - r}{\rho - r} R(m) - 1 < 0 \Rightarrow L(m) < \rho - r.$$

Since, in addition, Lemma A.1 implies that $\lim_{w_t \to \infty} L(w_t) = \rho - r$, there exists w_t large enough so that $L(w_t) > L(m) > 0$. These inequalities, together with $L'(m) < 0$, imply that $L(w_t)$ has a local minimum in (m, ∞). We can, therefore, choose $\xi \in (0, L(m))$, and $m_2 > m_1 > m$ such that $L(m_i) = \xi$, $L'(m_1) < 0$ and $L'(m_2) > 0$. Using (A.25), we can write (A.57) as

$$L'(w_t) = Q(w_t) \frac{R(w_t) - \gamma}{\gamma w_t} R(w_t) + [Q(w_t) - r] \frac{R(w_t)[1 - R(w_t)] + \frac{[\alpha w_t + R(w_t)]^2}{z} \left(\frac{Q(w_t) - r}{\rho - r} R(w_t) - 1 \right)}{w_t}. \tag{A.58}$$

Since

$$L(w_t) = \xi \Rightarrow Q(w_t) = \frac{\xi}{R(w_t)} + r, \tag{A.59}$$

we can substitute $Q(w_t)$ in $L'(w_t)$ to write the inequalities $L'(m_1) < 0$ and $L'(m_2) > 0$ as $\hat{L}_\xi(m_1) < 0$ and $\hat{L}_\xi(m_2) > 0$, respectively, where the function $\hat{L}_\xi(w_t)$ is defined by

$$\hat{L}_\xi(w_t) \equiv r \frac{R(w_t) - \gamma}{\gamma} + \left(\frac{1}{\gamma} - 1 \right) \xi - \frac{[\alpha w_t + R(w_t)]^2}{R(w_t)^2} \frac{\rho - r - \xi}{\rho - r} \xi. \tag{A.60}$$

Since $Q(w_t)$ is increasing, (A.59) implies $R(m_1) > R(m_2)$. Since $A(w_t)$ is decreasing (Lemma A.6), $\frac{\alpha w_t + R(w_t)}{R(w_t)} = \frac{\alpha}{A(w_t)} + 1$ is increasing. Since, in addition, $\xi \in (0, \rho - r)$, (A.60) implies $\hat{L}_\xi(m_1) > \hat{L}_\xi(m_2)$, a contradiction. Therefore, $L'(m) \geq 0$, which in turn implies that m is infinite.

Consider next the case $\gamma < K$. To show the result it suffices to show that $R'(w_t) > -\alpha$ for w_t close to zero, since we can then define the supremum m and proceed as in the case $K < \gamma < 1$. To show that $R'(w_t) > -\alpha$ for w_t close to zero, we study the behavior of $(Q(w_t), R(w_t))$. Since $\lim_{w_t \to 0} R(w_t) = K$, (A.25) implies

$$Q(w_t) \approx Q(\epsilon) \left(\frac{w_t}{\epsilon} \right)^{\frac{K - \gamma}{\gamma}} \tag{A.61}$$

for small ϵ and w_t. If $\frac{K-\gamma}{\gamma} > 1 \Leftrightarrow K > 2\gamma$, then $Q(w_t)$ is of order smaller than w_t, and we look for a differentiable $R(w_t)$. Using (A.26), $R(0) = K$ and L'Hospital's rule, we find

$$R'(0) = R'(0)(1-2K) - \frac{2[\alpha + R'(0)]K}{z}\left(\frac{rK}{\rho - r} + 1\right) - \frac{K^2}{z}\frac{rR'(0)}{\rho - r}$$

$$\Rightarrow R'(0) = -\frac{\frac{2\alpha K}{z}\left(\frac{rK}{\rho-r} + 1\right)}{2K + \frac{2K}{z}\left(\frac{rK}{\rho-r} + 1\right) + \frac{K^2}{z}\frac{r}{\rho-r}} > -\alpha.$$

If $\frac{K-\gamma}{\gamma} < 1 \Leftrightarrow K < 2\gamma$, then $Q(w_t)$ is of order larger than w_t, and we look for $R(w_t)$ that is of the form $K + gw^{\frac{K-\gamma}{\gamma}} + o\left(w^{\frac{K-\gamma}{\gamma}}\right)$. Substituting into (A.26), we find

$$\begin{aligned} g\frac{K-\gamma}{\gamma}w^{\frac{K-\gamma}{\gamma}} &= \left(K + gw^{\frac{K-\gamma}{\gamma}}\right)\left(1 - K - gw^{\frac{K-\gamma}{\gamma}}\right) \\ &\quad + \frac{\left(\alpha w_t + K + gw^{\frac{K-\gamma}{\gamma}}\right)^2}{z}\left(\frac{Q(\epsilon)\left(\frac{w_t}{\epsilon}\right)^{\frac{K-\gamma}{\gamma}} - r}{\rho - r}\left(K + gw^{\frac{K-\gamma}{\gamma}}\right) - 1\right). \end{aligned}$$

Identifying terms in $w^{\frac{K-\gamma}{\gamma}}$, we find

$$g\frac{K-\gamma}{\gamma} = g(1-2K) - \frac{2gK}{z}\left(\frac{rK}{\rho-r} + 1\right) + \frac{K^2}{z}\frac{Q(\epsilon)\left(\frac{1}{\epsilon}\right)^{\frac{K-\gamma}{\gamma}}K - rg}{\rho - r}$$

$$\Rightarrow g = \frac{\frac{K^3}{z}\frac{Q(\epsilon)\left(\frac{1}{\epsilon}\right)^{\frac{K-\gamma}{\gamma}}}{\rho-r}}{\frac{K-\gamma}{\gamma} - (1-2K) + \frac{2K}{z}\left(\frac{rK}{\rho-r} + 1\right) + \frac{K^2}{z}\frac{r}{\rho-r}} = \frac{\frac{K^2}{z}\frac{Q(\epsilon)\left(\frac{1}{\epsilon}\right)^{\frac{K-\gamma}{\gamma}}}{\rho-r}}{\frac{1}{\gamma} + \frac{K}{z}\frac{r}{\rho-r}},$$

where the last step follows from (3.23). Since $g > 0$, $R(w_t)$ is increasing for w_t close to zero, which means that $R'(w_t) \geq 0 > -\alpha$.

We next extend the proof to the solution constructed in Lemma A.3. Consider first the case $K < \gamma < 1$. Suppose, by contradiction, that the supremum m of the set of $\hat{w}_t$ such that $R'(w_t) > -\alpha$ for all $w_t \in [\epsilon, \hat{w}_t)$ is smaller than M_0. To derive a contradiction, it suffices to show that $L'(m) \geq 0$. Recall from Lemma A.5 that for any given w_t, the values of $Q(w_t)$ and $R(w_t)$ under the solution over $[\epsilon, M]$ converge to the values under the solution over $[\epsilon, \infty)$ when M goes to infinity. Since $m \leq M_0$ and M_0 is kept fixed when M grows large, $L(m)$ is close to its positive limit under the solution over $[\epsilon, \infty)$ for M large enough. Hence, $L(m) < \rho - r$. Moreover, same arguments as in Lemmas A.1 and A.5 imply that under the solution over $[\epsilon, \infty)$, $\lim_{w_t \to \infty} Q(w_t) = r + \frac{\rho - r}{\gamma}$ and $\lim_{w_t \to \infty} R(w_t) = \gamma$

(as is the case under the solution over $(0, \infty)$). Hence, there exists w_t large enough such that $L(w_t)$ is close to $\rho - r$ and hence larger than $L(m)$. We can then derive a contradiction, proceeding as above. Consider next the case $\gamma < K$. Since $R'(w_t) > -\alpha$ for small w_t under the solution over $(0, \infty)$, the same inequality holds under the solution over $[\epsilon, M]$ for ϵ small enough by continuity. We can then derive a contradiction, proceeding as above. ∎

Proof of Theorem 3.2: The comparative statics in the theorem can be stated equivalently in terms of $R(w_t)$: $R(w_t)$ is increasing in α and $u^\top \Sigma u$ if $\gamma < 1$, and is decreasing in α and $u^\top \Sigma u$ if $\gamma > 1$. We show the latter comparative statics on the solution constructed in Lemma A.3, which is over the finite interval $[\epsilon, M]$. We fix any $M_0 > \epsilon$, and show that the comparative statics hold for all $w_t \in (\epsilon, M_0)$, provided that ϵ and M are small and large enough, respectively. Since for any given $w_t \in (0, \infty)$, $(Q(w_t), R(w_t))$ for the solution over $(0, \infty)$ are obtained as limits of $(Q(w_t), R(w_t))$ for the solution over $[\epsilon, M]$ when ϵ goes to zero and M goes to infinity, the inequalities established for (ϵ, M_0) carry through to $(0, \infty)$.

The comparative statics with respect to $u^\top \Sigma u$ are equivalent to those with respect to z, and we show the latter. We start with the case $K < \gamma < 1$, and denote the solution to the system of (A.25) and (A.26) by the subscript z. Since $R_z(\epsilon) = \gamma$ for all z, differentiating (A.26) with respect to z, we find

$$\frac{\partial R'_z(\epsilon)}{\partial z}\epsilon = \frac{(\alpha\epsilon + \gamma)^2}{z}\left[-\frac{1}{z}\left(\frac{Q_z(\epsilon) - r}{\rho - r}\gamma - 1\right) + \frac{\frac{\partial Q_z(\epsilon)}{\partial z}}{\rho - r}\gamma\right]. \tag{A.62}$$

We will show that (i) $R_z(w_t) < R_{\hat{z}}(w_t)$ for $\hat{z}$ close to and larger than z and for all $w_t > \epsilon$ in a neighborhood of ϵ, and (ii) $R_z(w_t) \leq R_{\hat{z}}(w_t)$ for $\hat{z}$ close to and larger than z, for all $w_t \in (\epsilon, M_0)$, and for M large enough. These results will be shown by contradiction.

Suppose that $R_z(w_t) > R_{\hat{z}}(w_t)$ for $\hat{z}$ close to and larger than z and for all $w_t > \epsilon$ in a neighborhood of ϵ. Since $R_z(M) = R_{\hat{z}}(M) = \gamma$, the set of $w_t > \epsilon$ such that $R_z(w_t) = R_{\hat{z}}(w_t)$ is non-empty. The infimum m within that set is strictly larger than ϵ since $R_z(w_t) > R_{\hat{z}}(w_t)$ for all $w_t > \epsilon$ in a neighborhood of ϵ. Since $R_{\hat{z}}(w_t)$ crosses $R_z(w_t)$ from above at ϵ, $\frac{\partial R'_z(\epsilon)}{\partial z} \leq 0$, and since it crosses $R_z(w_t)$ from below at m, $\frac{\partial R'_z(m)}{\partial z} \geq 0$.[14] The counterpart of (A.62) at m is

$$\frac{\partial R'_z(m)}{\partial z} m = \frac{[\alpha m + R_z(m)]^2}{z} \left[-\frac{1}{z} \left(\frac{Q_z(m) - r}{\rho - r} R_z(m) - 1 \right) + \frac{\frac{\partial Q_z(m)}{\partial z}}{\rho - r} R_z(m) \right] \tag{A.63}$$

$$= \frac{[\alpha m + R_z(m)]^2}{z} \left[-\frac{1}{z} \left(\frac{Q_z(m) - r}{\rho - r} R_z(m) - 1 \right) \right.$$

$$\left. + \frac{\frac{\partial Q_z(\epsilon)}{\partial z} \exp\left(\int_\epsilon^m \frac{R_z(w_t) - \gamma}{\gamma w_t} dw_t \right) + Q_z(\epsilon) \frac{\partial \exp\left(\int_\epsilon^m \frac{R_z(w_t)-\gamma}{\gamma w_t} dw_t \right)}{\partial z}}{\rho - r} R_z(m) \right], \tag{A.64}$$

where the second step follows because integrating (A.25) from ϵ to m yields

$$Q_z(m) = Q_z(\epsilon) \exp\left(\int_\epsilon^m \frac{R_z(w_t) - \gamma}{\gamma w_t} dw_t \right). \tag{A.65}$$

Equations (A.62) and $\frac{\partial R'_z(\epsilon)}{\partial z} \leq 0$ imply

$$-\frac{1}{z} \left(\frac{Q_z(\epsilon) - r}{\rho - r} \gamma - 1 \right) + \frac{\frac{\partial Q_z(\epsilon)}{\partial z}}{\rho - r} \gamma \leq 0$$

$$\Rightarrow \left[-\frac{1}{z} \left(\frac{Q_z(\epsilon) - r}{\rho - r} \gamma - 1 \right) + \frac{\frac{\partial Q_z(\epsilon)}{\partial z}}{\rho - r} \gamma \right] \frac{Q_z(m) R_z(m)}{Q_z(\epsilon)\gamma} \leq 0$$

$$\Rightarrow -\frac{1}{z} \left(\frac{Q_z(m) - r}{\rho - r} R_z(m) - 1 \right) + \frac{\frac{\partial Q_z(\epsilon)}{\partial z}}{\rho - r} \frac{Q_z(m)}{Q_z(\epsilon)} R_z(m)$$

$$+ \frac{1}{z} \left(\frac{Q_z(m) R_z(m)}{Q_z(\epsilon)\gamma} - 1 \right) + \frac{r}{z(\rho - r)} \left(\frac{Q_z(m)}{Q_z(\epsilon)} - 1 \right) R_z(m) \leq 0$$

$$\Rightarrow -\frac{1}{z} \left(\frac{Q_z(m) - r}{\rho - r} R_z(m) - 1 \right) + \frac{\frac{\partial Q_z(\epsilon)}{\partial z}}{\rho - r} \frac{Q_z(m)}{Q_z(\epsilon)} R_z(m) < 0$$

$$\Rightarrow -\frac{1}{z} \left(\frac{Q_z(m) - r}{\rho - r} R_z(m) - 1 \right)$$

$$+ \frac{\frac{\partial Q_z(\epsilon)}{\partial z} \exp\left(\int_\epsilon^m \frac{R_z(w_t) - \gamma}{\gamma w_t} dw_t \right) + Q_z(\epsilon) \frac{\partial \exp\left(\int_\epsilon^m \frac{R_z(w_t)-\gamma}{\gamma w_t} dw_t \right)}{\partial z}}{\rho - r} R_z(m) < 0$$

$$\Rightarrow \frac{\partial R'_z(m)}{\partial z} < 0,$$

[14] For expositional simplicity, we are treating m as independent of $\hat{z}$ for $\hat{z}$ close to z. When m is a function $m(\hat{z})$ of $\hat{z}$, the property $\frac{\partial R'_z(m)}{\partial z} \geq 0$ holds for a non-empty set of m, e.g., for $\liminf_{\hat{z} \to z} m(\hat{z})$ and $\limsup_{\hat{z} \to z} m(\hat{z})$. Our argument requires choosing one such m.

where the fourth step follows because $Q_z(w_t)$ is increasing and $R(w_t) > \gamma$ for all $w_t \in (\epsilon, M)$, the fifth step follows from (A.65) and because $R_z(w_t) > R_{\hat{z}}(w_t)$ for all $w_t \in (\epsilon, m)$ implies $\frac{\partial \exp\left(\int_\epsilon^m \frac{R_z(w_t)-\gamma}{\gamma w_t} dw_t\right)}{\partial z} \leq 0$, and the sixth step follows from (A.64). This contradicts $\frac{\partial R_z'(m)}{\partial z} \geq 0$. Hence, it is not possible that $R_z(w_t) > R_{\hat{z}}(w_t)$ for all $w_t > \epsilon$ in a neighborhood of ϵ. This implies that $R_z(w_t) < R_{\hat{z}}(w_t)$ for all $w_t > \epsilon$ in a neighborhood of ϵ.[15]

Consider next the infimum $m > \epsilon$ within the set of $w_t > \epsilon$ such that $R_z(w_t) = R_{\hat{z}}(w_t)$ and $R_z(\hat{w}_t) > R_{\hat{z}}(\hat{w}_t)$ for all $\hat{w}_t > w_t$ in a neighborhood of w_t. This set is non-empty because it includes M. Our intended comparative statics result will follow if $m \geq M_0$ since in that case $R_z(w_t) \leq R_{\hat{z}}(w_t)$ for all $w_t \in (\epsilon, M_0)$. Suppose that $m < M_0$. Since $R_z(w_t) > R_{\hat{z}}(w_t)$ for all $w_t > m$ in a neighborhood of m, we can consider the infimum $\hat{m} > m$ within the non-empty set of $w_t > m$ such that $R_z(w_t) = R_{\hat{z}}(w_t)$. Since $R_{\hat{z}}(w_t)$ crosses $R_z(w_t)$ from above at m, $\frac{\partial R_z'(m)}{\partial z} \leq 0$, and since it crosses $R_z(w_t)$ from below at $\hat{m}$, $\frac{\partial R_z'(\hat{m})}{\partial z} \geq 0$.

Equations (A.63) and $\frac{\partial R_z'(m)}{\partial z} \leq 0$ imply

$$-\frac{1}{z}\left(\frac{Q_z(m)-r}{\rho-r}R_z(m)-1\right)+\frac{\frac{\partial Q_z(m)}{\partial z}}{\rho-r}R_z(m) \leq 0$$

$$\Rightarrow -\frac{1}{zQ_z(m)R_z(m)}\left(\frac{Q_z(m)-r}{\rho-r}R_z(m)-1\right) \leq -\frac{\frac{\partial Q_z(m)}{\partial z}}{Q_z(m)(\rho-r)}. \tag{A.66}$$

The counterpart of (A.64) written between m and $\hat{m}$ instead of between ϵ and m, $\frac{\partial R_z'(\hat{m})}{\partial z} \geq 0$ and $\frac{\partial \exp\left(\int_m^{\hat{m}} \frac{R_z(w_t)-\gamma}{\gamma w_t} dw_t\right)}{\partial z} \leq 0$ (which follows from $R_z(w_t) > R_{\hat{z}}(w_t)$ for all $w_t \in (m, \hat{m})$) imply

$$-\frac{1}{z}\left(\frac{Q_z(\hat{m})-r}{\rho-r}R_z(\hat{m})-1\right)+\frac{\frac{\partial Q_z(m)}{\partial z}\exp\left(\int_m^{\hat{m}}\frac{R_z(w_t)-\gamma}{\gamma w_t}dw_t\right)}{\rho-r}R_z(\hat{m}) \geq 0$$

$$\Rightarrow -\frac{1}{z}\left(\frac{Q_z(\hat{m})-r}{\rho-r}R_z(\hat{m})-1\right)+\frac{\frac{\partial Q_z(m)}{\partial z}\frac{Q_z(\hat{m})}{Q_z(m)}}{\rho-r}R_z(\hat{m}) \geq 0$$

$$\Rightarrow -\frac{1}{zQ_z(\hat{m})R_z(\hat{m})}\left(\frac{Q_z(\hat{m})-r}{\rho-r}R_z(\hat{m})-1\right) \geq -\frac{\frac{\partial Q_z(m)}{\partial z}}{Q_z(m)(\rho-r)}, \tag{A.67}$$

where the second step follows from the counterpart of (A.65) written between m and $\hat{m}$ and the third step follows by dividing both sides by $Q_z(m)R_z(m)$.

[15] The implication follows by showing that not all the derivatives $\frac{\partial R_z^{(n)}(\epsilon)}{\partial z}$ are equal to zero.

Equations (A.66) and (A.67) imply $N_z(m) \leq N_z(\hat{m})$, where

$$N_z(w_t) \equiv -\frac{1}{zQ_z(w_t)R_z(w_t)}\left(\frac{Q_z(w_t)-r}{\rho-r}R_z(w_t)-1\right).$$

Equation (A.26) implies $N_z(w_t) > 0$ for all w_t for which $R_z'(w_t) \leq 0$. The same inequality holds for all w_t for which $R_z(w_t) > 0$: this follows because $Q_z(w_t)$ is increasing and because for all w_t such that $R_z(w_t) > 0$, there exists $\hat{w}_t > w_t$ such that $R_z(w_t) = R_z(\hat{w}_t)$ and $R_z'(\hat{w}_t) \leq 0$. Consider now M large enough, and recall from Lemma A.5 that for any given w_t, the values of $Q_z(w_t)$ and $R_z(w_t)$ under the solution to the system of (A.25) and (A.26) over $[\epsilon, M]$ converge to the values under the solution over $[\epsilon, \infty)$ when M goes to infinity. Since $m < M_0$ and M_0 is kept fixed when M grows large, $N_z(m)$ is close to its positive limit under the solution over $[\epsilon, \infty)$ for M large enough. Moreover, same arguments as in Lemmas A.1 and A.5 imply that under the solution over $[\epsilon, \infty)$, $\lim_{w_t \to \infty} Q_z(w_t) = r + \frac{\rho-r}{\gamma}$ (as is the case under the solution over $(0, \infty)$). Picking $\overline{m}$ such that $Q_z(\overline{m})$ is close to $r + \frac{\rho-r}{\gamma}$ for M large enough, and noting that $Q_z(w_t)$ is increasing and $R_z(w_t) > \gamma$, we can bound $N_z(w_t)$ from above by a small positive constant for all $w_t \geq \overline{m}$ and for M large enough. Hence, for all $w_t \geq \overline{m}$ and for M large enough, $N_z(w_t)$ is smaller than $N_z(m)$ minus a positive constant.

If the inequality $N_z(m) \leq N_z(\hat{m})$ is strict, then we can choose $\xi \in (N_z(m), N_z(\hat{m}))$, $m_1 \in (m, \hat{m})$ and $m_2 \in (\hat{m}, M)$, such that $N_z(m_i) = \xi$ for $i = 1, 2$, $G'(m_1) \geq 0$ and $G'(m_2) \leq 0$. Values $\xi > 0$ and $m_1 < m_2$ such that $N_z(m_i) = \xi$ for $i = 1, 2$, $G'(m_1) \geq 0$ and $G'(m_2) \leq 0$ can also be chosen if $N_z(m) = N_z(\hat{m})$. Indeed, if $\min_{w_t \in [m, \hat{m}]} N_z(w_t) < N_z(m) = N_z(\hat{m})$, then we choose ξ close to and smaller than $N_z(m) = N_z(\hat{m})$, $m_1 \in (m, \hat{m})$ and $m_2 \in (\hat{m}, M)$. If $N_z(w_t) = N_z(m) = N_z(\hat{m})$ for all $w_t \in (m, \hat{m})$, then we choose $\xi = N_z(m) = N_z(\hat{m})$ and any $m_1 < m_2$ in $(m, \hat{m})$. If, finally, $\max_{w_t \in [m, \hat{m}]} N_z(w_t) > N_z(m) = N_z(\hat{m})$, then we replace $\hat{m}$ by the maximizer of $N_z(w_t)$ in $[m, \hat{m}]$, and proceed as in the case where the inequality is strict.

To compute $N_z'(w_t)$, we write $N_z(w_t)$ as

$$N_z(w_t) = -\frac{1}{z(\rho-r)} + \frac{r}{z(\rho-r)Q_z(w_t)} + \frac{1}{zQ_z(w_t)R_z(w_t)}.$$

Equations (A.25) and (A.26) imply

$$N_z'(w_t) = -\frac{\frac{r}{\rho-r}\frac{R_z(w_t)-\gamma}{\gamma} + \frac{1}{\gamma} - 1 + \frac{[\alpha w_t + R_z(w_t)]^2}{zR_z(w_t)^2}\left(\frac{Q_z(w_t)-r}{\rho-r}R_z(w_t)-1\right)}{zw_tQ_z(w_t)}.$$

Since

$$N_z(w_t) = \xi \Rightarrow Q_z(w_t) = \frac{1}{1+z(\rho-r)\xi}\left(\frac{\rho-r}{R_z(w_t)}+r\right), \tag{A.68}$$

we can substitute $Q_z(w_t)$ in $N_z'(w_t)$ to write the inequalities $N_z'(m_1) \geq 0$ and $N_z'(m_2) \leq 0$ as $\hat{N}_{z\xi}(m_1) \leq 0$ and $\hat{N}_{z\xi}(m_2) \geq 0$, respectively, where the function $\hat{N}_{z\xi}(w_t)$ is defined by

$$\hat{N}_{z\xi}(w_t) \equiv \frac{r}{\rho-r}\frac{R_z(w_t)-\gamma}{\gamma}+\frac{1}{\gamma}-1-\frac{[\alpha w_t+R_z(w_t)]^2}{R_z(w_t)^2}\frac{rR_z(w_t)+\rho-r}{1+z(\rho-r)\xi}\xi. \tag{A.69}$$

We will show that $\hat{N}_{z\xi}(m_1) > \hat{N}_{z\xi}(m_2)$, which will yield a contradiction because $\hat{N}_{z\xi}(m_1) \leq 0$ and $\hat{N}_{z\xi}(m_2) \geq 0$. As a first step to show that result, we will show that

$$\alpha m_1 + R_z(m_1) < \alpha m_2 + R_z(m_2). \tag{A.70}$$

If $m_2 > m_1 + \frac{1-\gamma}{\alpha}$, then (A.70) follows from $R_z(w_t) \in (\gamma, 1)$ for all $w_t \in (\epsilon, M)$. If $m_2 < m_1 + \frac{1-\gamma}{\alpha}$, then (A.70) follows from Lemma A.8 by noting that $m_1 < \hat{m} < \overline{m}$ and setting $M_0 = \overline{m} + \frac{1-\gamma}{\alpha}$. Using (A.69), (A.70) and $\xi > 0$, we find

$$\hat{N}_{z\xi}(m_1) > \frac{r}{\rho-r}\frac{R_z(m_1)-\gamma}{\gamma}+\frac{1}{\gamma}-1-[\alpha m_2+R_z(m_2)]\frac{\alpha m_1+R_z(m_1)}{R_z(m_1)}\frac{rR_z(w_t)+\rho-r}{R_z(m_1)[1+z(\rho-r)\xi]}\xi. \tag{A.71}$$

Since $Q_z(w_t)$ is increasing, (A.68) implies $R_z(m_1) > R_z(m_2)$. Since $A(w_t)$ is decreasing (Lemma A.6), $\frac{\alpha w_t+R_z(w_t)}{R_z(w_t)} = \frac{\alpha}{A(w_t)}+1$ is increasing. Since, in addition, $\frac{rR_z(w_t)+\rho-r}{R_z(w_t)} = r+\frac{\rho-r}{R_z(w_t)}$ is decreasing in $R_z(w_t)$, and $\xi > 0$, (A.71) implies $\hat{N}_{z\xi}(m_1) > \hat{N}_{z\xi}(m_2)$.

Consider next the case $\gamma < K$. Since $R_z(\epsilon)$ is increasing in z, $R_z(w_t) < R_{\hat{z}}(w_t)$ for $\hat{z}$ close to and larger than z and for all $w_t > \epsilon$ in a neighborhood of ϵ. Therefore, we are left to show result (ii) of the case $K < \gamma < 1$, namely, $R_z(w_t) \leq R_{\hat{z}}(w_t)$ for $\hat{z}$ close to and larger than z, for all $w_t \in (\epsilon, M_0)$, and for M large enough. We define m as in that case, suppose that $m < M_0$, and define $\hat{m}$. We then follow the same arguments to derive a contradiction.

Consider finally the case $\gamma > 1$. We will show, proceeding by contradiction, that (i) $R_z(w_t) > R_{\hat{z}}(w_t)$ for $\hat{z}$ close to and larger than z and for all $w_t > \epsilon$ in a neighborhood of ϵ, and (ii) $R_z(w_t) \geq R_{\hat{z}}(w_t)$ for $\hat{z}$ close to and larger than z, for all $w_t \in (\epsilon, M_0)$, and for M large enough.

Suppose that $R_z(w_t) < R_{\hat{z}}(w_t)$ for $\hat{z}$ close to and larger than z and for all $w_t > \epsilon$ in a neighborhood of ϵ. Consider the infimum $m > \epsilon$ within the non-empty set of $w_t > \epsilon$ such that $R_z(w_t) = R_{\hat{z}}(w_t)$. Since $R_{\hat{z}}(w_t)$ crosses $R_z(w_t)$ from below at ϵ, $\frac{\partial R_z'(\epsilon)}{\partial z} \geq 0$, and since it crosses $R_z(w_t)$ from above at m, $\frac{\partial R_z'(m)}{\partial z} \leq 0$. Equations (A.62) and $\frac{\partial R_z'(\epsilon)}{\partial z} \geq 0$ imply

$$
\begin{aligned}
&-\frac{1}{z}\left(\frac{Q_z(\epsilon)-r}{\rho-r}\gamma - 1\right) + \frac{\frac{\partial Q_z(\epsilon)}{\partial z}}{\rho - r}\gamma \geq 0 \\
&\Rightarrow -\frac{1}{z}\left(\frac{Q_z(m)-r}{\rho-r}R_z(m) - 1\right) + \frac{\frac{\partial Q_z(\epsilon)}{\partial z}}{\rho - r}\frac{Q_z(m)}{Q_z(\epsilon)}R_z(m) \\
&\quad + \frac{1}{z}\left(\frac{Q_z(m)R_z(m)}{Q_z(\epsilon)\gamma} - 1\right) + \frac{r}{z(\rho-r)}\left(\frac{Q_z(m)}{Q_z(\epsilon)} - 1\right)R_z(m) \geq 0 \\
&\Rightarrow -\frac{1}{z}\left(\frac{Q_z(m)-r}{\rho-r}R_z(m) - 1\right) + \frac{\frac{\partial Q_z(\epsilon)}{\partial z}}{\rho - r}\frac{Q_z(m)}{Q_z(\epsilon)}R_z(m) > 0 \\
&\Rightarrow -\frac{1}{z}\left(\frac{Q_z(m)-r}{\rho-r}R_z(m) - 1\right) \\
&\quad + \frac{\frac{\partial Q_z(\epsilon)}{\partial z}\exp\left(\int_\epsilon^m \frac{R_z(w_t)-\gamma}{\gamma w_t}dw_t\right) + Q_z(\epsilon)\frac{\partial \exp\left(\int_\epsilon^m \frac{R_z(w_t)-\gamma}{\gamma w_t}dw_t\right)}{\partial z}}{\rho - r}R_z(m) > 0 \\
&\Rightarrow \frac{\partial R_z'(m)}{\partial z} > 0,
\end{aligned}
$$

where the third step follows because $Q_z(w_t)$ is decreasing and $R(w_t) < \gamma$ for all $w_t \in (\epsilon, M)$, the fourth step follows from (A.65) and because $R_z(w_t) < R_{\hat{z}}(w_t)$ for all $w_t \in (\epsilon, m)$ implies $\frac{\partial \exp\left(\int_\epsilon^m \frac{R_z(w_t)-\gamma}{\gamma w_t}dw_t\right)}{\partial z} \geq 0$, and the fifth step follows from (A.64). This contradicts $\frac{\partial R_z'(m)}{\partial z} \leq 0$. Hence, $R_z(w_t) > R_{\hat{z}}(w_t)$ for all $w_t > \epsilon$ in a neighborhood of ϵ.

Consider next the infimum $m > \epsilon$ within the non-empty set of $w_t > \epsilon$ such that $R_z(w_t) = R_{\hat{z}}(w_t)$ and $R_z(\hat{w}_t) < R_{\hat{z}}(\hat{w}_t)$ for all $\hat{w}_t > w_t$ in a neighborhood of w_t. Our intended comparative statics result will follow if $m \geq M_0$ since in that case $R_z(w_t) \geq R_{\hat{z}}(w_t)$ for all $w_t \in (\epsilon, M_0)$. Suppose that $m < M_0$. Since $R_z(w_t) > R_{\hat{z}}(w_t)$ for all $w_t > m$ in a neighborhood of m, we can consider the infimum $\hat{m} > m$ within the non-empty set of $w_t > m$ such that $R_z(w_t) = R_{\hat{z}}(w_t)$. Since $R_{\hat{z}}(w_t)$ crosses $R_z(w_t)$ from below at m, $\frac{\partial R_z'(m)}{\partial z} \geq 0$, and since it crosses $R_z(w_t)$ from above at $\hat{m}$, $\frac{\partial R_z'(\hat{m})}{\partial z} \leq 0$. Proceeding as in the case $K < \gamma < 1$, we can then show that $N_z(m) \geq N_z(\hat{m})$, $-\frac{1}{z(\rho-r)} < N_z(w_t) < 0$ for all $w_t \in [\epsilon, M]$, and $N_z(w_t)$ exceeds $N_z(m)$ plus a positive constant for all

w_t larger than a fixed $\overline{m}$ and for M large enough. We can then choose $\xi \in (-\frac{1}{z(\rho-r)}, 0)$ and $m_1 < m_2$ such that $N_z(m_i) = \xi$ for $i = 1, 2$, $N_z'(m_1) \le 0$ and $N_z'(m_2) \ge 0$. The latter two inequalities can be written as $\hat{N}_{z\xi}(m_1) \ge 0$ and $\hat{N}_{z\xi}(m_2) \le 0$. We will show that $\hat{N}_{z\xi}(m_1) < \hat{N}_{z\xi}(m_2)$, which will yield a contradiction because $\hat{N}_{z\xi}(m_1) \ge 0$ and $\hat{N}_{z\xi}(m_2) \le 0$. As a first step to show that result, we will show that

$$A_z(m_1) > A_z(m_2). \tag{A.72}$$

If $m_2 > \gamma m_1$, then (A.72) follows from $R_z(w_t) \in (1, \gamma)$ for all $w_t \in (\epsilon, M)$. If $m_2 > \gamma m_1$, then (A.72) follows from Lemma A.6 by noting that $m_1 < \hat{m} < \overline{m}$ and setting $M_0 = \gamma m_1$. Using $\frac{\alpha m_1 + R_z(m_1)}{R_z(m_1)} < \frac{\alpha m_2 + R_z(m_2)}{R_z(m_2)}$, which is implied from (A.72), and $\xi \in (-\frac{1}{z(\rho-r)}, 0)$, we find

$$\hat{N}_{z\xi}(m_1) < \frac{r}{\rho - r}\frac{R_z(m_1) - \gamma}{\gamma} + \frac{1}{\gamma} - 1 - \frac{[\alpha m_2 + R_z(m_2)]^2}{R_z(m_2)^2}\frac{rR_z(w_t) + \rho - r}{1 + z(\rho - r)\xi}\xi. \tag{A.73}$$

Since $Q_z(w_t)$ is decreasing, (A.68) implies $R_z(m_1) < R_z(m_2)$. Combining with (A.71) and $\xi \in (-\frac{1}{z(\rho-r)}, 0)$, we find $\hat{N}_{z\xi}(m_1) < \hat{N}_{z\xi}(m_2)$.

We next derive comparative statics with respect to α. We start with the case $K < \gamma < 1$, and denote the solution to the system of (A.25) and (A.26) by the subscript α. Since $R_\alpha(\epsilon) = \gamma$ for all α, differentiating (A.25) with respect to α and noting that z depends on α through (3.20), we find

$$\frac{\partial R_a'(\epsilon)}{\partial z}\epsilon = \frac{(\alpha\epsilon + \gamma)\gamma}{z}\left[-\frac{2}{\alpha}\left(\frac{Q_\alpha(\epsilon) - r}{\rho - r}\gamma - 1\right) + (\alpha\epsilon + \gamma)\frac{\frac{\partial Q_\alpha(\epsilon)}{\partial z}}{\rho - r}\right]. \tag{A.74}$$

We will show that (i) $R_\alpha(w_t) < R_{\hat{\alpha}}(w_t)$ for $\hat{\alpha}$ close to and larger than α and for all $w_t > \epsilon$ in a neighborhood of ϵ, and (ii) $R_\alpha(w_t) \le R_{\hat{\alpha}}(w_t)$ for $\hat{\alpha}$ close to and larger than α, for all $w_t \in (\epsilon, M_0)$, and for M large enough. These results will be shown by contradiction, using similar arguments as for the comparative statics with respect to z.

Suppose that $R_\alpha(w_t) > R_{\hat{\alpha}}(w_t)$ for $\hat{\alpha}$ close to and larger than α and for all $w_t > \epsilon$ in a neighborhood of ϵ. Consider the infimum $m > \epsilon$ within the non-empty set of $w_t > \epsilon$ such that $R_\alpha(w_t) = R_{\hat{\alpha}}(w_t)$. Since $R_{\hat{\alpha}}(w_t)$ crosses $R_\alpha(w_t)$ from above at ϵ, $\frac{\partial R_\alpha'(\epsilon)}{\partial z} \le 0$, and since it crosses

$R_\alpha(w_t)$ from below at m, $\frac{\partial R'_\alpha(m)}{\partial z} \geq 0$. The counterpart of (A.74) at m is

$$\frac{\partial R_z'(m)}{\partial z}m = \frac{[\alpha m + R_\alpha(m)]R_\alpha(m)}{z}\left[-\frac{2}{\alpha}\left(\frac{Q_\alpha(m)-r}{\rho-r}R_\alpha(m)-1\right)+[\alpha m + R_\alpha(m)]\frac{\frac{\partial Q_\alpha(m)}{\partial z}}{\rho-r}\right] \tag{A.75}$$

$$= \frac{[\alpha m + R_\alpha(m)]R_\alpha(m)}{z}\left[-\frac{2}{\alpha}\left(\frac{Q_\alpha(m)-r}{\rho-r}R_\alpha(m)-1\right)\right.$$
$$\left. +[\alpha m + R_\alpha(m)]\frac{\frac{\partial Q_\alpha(\epsilon)}{\partial \alpha}\exp\left(\int_\epsilon^m \frac{R_\alpha(w_t)-\gamma}{\gamma w_t}dw_t\right)+Q_\alpha(\epsilon)\frac{\partial \exp\left(\int_\epsilon^m \frac{R_\alpha(w_t)-\gamma}{\gamma w_t}dw_t\right)}{\partial\alpha}}{\rho-r}\right]. \tag{A.76}$$

Equations (A.74) and $\frac{\partial R'_\alpha(\epsilon)}{\partial \alpha} \leq 0$ imply

$$-\frac{2}{\alpha}\left(\frac{Q_\alpha(\epsilon)-r}{\rho-r}\gamma-1\right)+(\alpha\epsilon+\gamma)\frac{\frac{\partial Q_\alpha(\epsilon)}{\partial z}}{\rho-r}\leq 0$$

$$\Rightarrow\left[-\frac{2}{\alpha}\left(\frac{Q_\alpha(\epsilon)-r}{\rho-r}\gamma-1\right)+(\alpha\epsilon+\gamma)\frac{\frac{\partial Q_\alpha(\epsilon)}{\partial z}}{\rho-r}\right]\frac{Q_\alpha(m)[\alpha m+R_\alpha(m)]}{Q_\alpha(\epsilon)(\alpha\epsilon+\gamma)}\leq 0$$

$$\Rightarrow -\frac{2}{\alpha}\left(\frac{Q_\alpha(m)-r}{\rho-r}R_\alpha(m)-1\right)+[\alpha m+R_\alpha(m)]\frac{\frac{\partial Q_\alpha(\epsilon)}{\partial \alpha}}{\rho-r}\frac{Q_\alpha(m)}{Q_\alpha(\epsilon)}$$
$$+\frac{2}{\alpha}\left(\frac{Q_\alpha(m)-r}{\rho-r}R_\alpha(m)-1\right)-\frac{2}{\alpha}\left(\frac{Q_\alpha(\epsilon)-r}{\rho-r}\gamma-1\right)\frac{Q_\alpha(m)[\alpha m+R_\alpha(m)]}{Q_\alpha(\epsilon)(\alpha\epsilon+\gamma)}\leq 0. \tag{A.77}$$

We will show that the term in the last line of (A.77) is positive. Since $A_\alpha(w_t)$ is decreasing (Lemma A.6),

$$\frac{\frac{\alpha}{A_\alpha(m)}+1}{\frac{\alpha}{A_\alpha(\epsilon)}+1}=\frac{[\alpha m+R_\alpha(m)]\gamma}{(\alpha\epsilon+\gamma)R_\alpha(m)}>1\Rightarrow\frac{\alpha m+R_\alpha(m)}{\alpha\epsilon+\gamma}>\frac{R_\alpha(m)}{\gamma}.$$

Since, in addition, $\frac{Q_\alpha(\epsilon)-r}{\rho-r}\gamma-1<0$, the term in the last line of (A.77) is larger than

$$\frac{2}{\alpha}\left(\frac{Q_\alpha(m)-r}{\rho-r}R_\alpha(m)-1\right)-\frac{2}{\alpha}\left(\frac{Q_\alpha(\epsilon)-r}{\rho-r}\gamma-1\right)\frac{Q_\alpha(m)R_\alpha(m)}{Q_\alpha(\epsilon)\gamma}$$
$$=\frac{2}{\alpha}\left(\frac{Q_\alpha(m)R_\alpha(m)}{Q_\alpha(\epsilon)\gamma}-1\right)+\frac{2r}{\alpha(\rho-r)}\left(\frac{Q_\alpha(m)}{Q_\alpha(\epsilon)}-1\right)R_\alpha(m),$$

which is positive because $Q_z(w_t)$ is increasing and $R(w_t) > \gamma$ for all $w_t \in (\epsilon, M)$. Hence, (A.77) implies

$$
\begin{aligned}
&-\frac{2}{\alpha}\left(\frac{Q_\alpha(m)-r}{\rho-r}R_\alpha(m)-1\right)+[\alpha m+R_\alpha(m)]\frac{\frac{\partial Q_\alpha(\epsilon)}{\partial\alpha}}{\rho-r}\frac{Q_\alpha(m)}{Q_\alpha(\epsilon)}<0\\
&\Rightarrow -\frac{2}{\alpha}\left(\frac{Q_\alpha(m)-r}{\rho-r}R_\alpha(m)-1\right)\\
&\quad+[\alpha m+R_\alpha(m)]\frac{\frac{\partial Q_\alpha(\epsilon)}{\partial\alpha}\exp\left(\int_\epsilon^m\frac{R_\alpha(w_t)-\gamma}{\gamma w_t}dw_t\right)+Q_\alpha(\epsilon)\frac{\partial\exp\left(\int_\epsilon^m\frac{R_\alpha(w_t)-\gamma}{\gamma w_t}dw_t\right)}{\partial\alpha}}{\rho-r}<0\\
&\Rightarrow \frac{\partial R'_\alpha(m)}{\partial\alpha}<0,
\end{aligned}
$$

where the second step follows from (A.65) (with subscript α rather than z) and because $R_\alpha(w_t) > R_{\hat{\alpha}}(w_t)$ for all $w_t \in (\epsilon, m)$ implies $\frac{\partial\exp\left(\int_\epsilon^m\frac{R_\alpha(w_t)-\gamma}{\gamma w_t}dw_t\right)}{\partial\alpha} \leq 0$, and the third step follows from (A.64). This contradicts $\frac{\partial R'_\alpha(m)}{\partial\alpha} \geq 0$. Hence, it is not possible that $R_\alpha(w_t) > R_{\hat{\alpha}}(w_t)$ for all $w_t > \epsilon$ in a neighborhood of ϵ. Instead, $R_\alpha(w_t) < R_{\hat{\alpha}}(w_t)$ for all $w_t > \epsilon$ in a neighborhood of ϵ.

Consider next the infimum $m > \epsilon$ within the non-empty set of $w_t > \epsilon$ such that $R_\alpha(w_t) = R_{\hat{\alpha}}(w_t)$ and $R_\alpha(\hat{w}_t) > R_{\hat{\alpha}}(\hat{w}_t)$ for all $\hat{w}_t > w_t$ in a neighborhood of w_t. Our intended comparative statics result will follow if $m \geq M_0$. Suppose that $m < M_0$. Since $R_\alpha(w_t) > R_{\hat{\alpha}}(w_t)$ for all $w_t > m$ in a neighborhood of m, we can consider the infimum $\hat{m} > m$ within the non-empty set of $w_t > m$ such that $R_\alpha(w_t) = R_{\hat{\alpha}}(w_t)$. Since $R_{\hat{\alpha}}(w_t)$ crosses $R_\alpha(w_t)$ from above at m, $\frac{\partial R'_\alpha(m)}{\partial\alpha} \leq 0$, and since it crosses $R_z(w_t)$ from below at $\hat{m}$, $\frac{\partial R'_\alpha(\hat{m})}{\partial\alpha} \geq 0$.

Proceeding as for the comparative statics with respect to z, we can show that $N_\alpha(m) \leq N_\alpha(\hat{m})$, where

$$
N_\alpha(w_t) \equiv -\frac{2}{\alpha Q_\alpha(w_t)[\alpha m+R_\alpha(w_t)]}\left(\frac{Q_\alpha(w_t)-r}{\rho-r}R_\alpha(w_t)-1\right),
$$

and that $N_\alpha(w_t) > 0$ for all $w_t \in (\epsilon, M)$. Since

$$
\frac{N_\alpha(m)}{N_\alpha(\hat{m})} = \frac{N_z(m)}{N_z(\hat{m})}\frac{[\alpha\hat{m}+R_\alpha(\hat{m})]R_\alpha(m)}{[\alpha m+R_\alpha(m)]R_\alpha(\hat{m})},
$$

and

$$
\frac{[\alpha\hat{m}+R_\alpha(\hat{m})]R_\alpha(m)}{[\alpha m+R_\alpha(m)]R_\alpha(\hat{m})} = \frac{\frac{\alpha}{A_\alpha(\hat{m})}+1}{\frac{\alpha}{A_\alpha(m)}+1} > 1
$$

because $A_\alpha(w_t)$ is decreasing (Lemma A.6), $N_\alpha(m) \leq N_\alpha(\hat{m})$ implies $N_z(m) < N_z(\hat{m})$. Since, in addition, $N_z(w_t)$ is smaller than $N_z(m)$ minus a positive constant for all w_t larger than a fixed $\overline{m}$ and for M large enough, the arguments used to establish the comparative statics with respect to z yield a contradiction. The comparative statics with respect to α in the cases $\gamma < K$ and $\gamma > 1$ follow by similarly adapting the arguments used to establish the comparative statics with respect to z. ∎

Proof of Corollary 3.2: The first statement in part (i) follows from (3.18), $F(w_t) = 0$, and $A(w_t)$ being decreasing in α when $\gamma > 1$ and independent of α when $\gamma = 1$. Since for $\gamma < K$ and small w_t, $A(w_t) \approx \frac{K}{w_t}$, the second statement in part (i) follows from (3.18) and $F(w_t) = 0$ if $\frac{\alpha}{K}$ is decreasing in α. Differentiating $\frac{\alpha}{K}$ with respect to α, we find

$$\frac{\partial \frac{\alpha}{K}}{\partial \alpha} = \frac{K - \alpha \frac{\partial K}{\partial \alpha}}{K^2} = \frac{K - \alpha \left(\frac{\frac{2K}{\alpha z}\left(\frac{rK}{\rho - r} + 1\right)}{1 + \frac{1}{z}\left(\frac{rK}{\rho - r} + 1\right) + \frac{rK}{z(\rho - r)}} \right)}{K^2} = \frac{K\left(1 - \frac{1}{z}\right)}{K^2\left[1 + \frac{1}{z}\left(\frac{rK}{\rho - r} + 1\right) + \frac{rK}{z(\rho - r)}\right]},$$

where the second step follows by differentiating implicitly K with respect to α using (3.23). Therefore, $\frac{\alpha}{K}$ is decreasing in α if $z < 1$. The third statement in part (i) follows from (3.19), $F(w_t) = 0$, and $A(w_t)$ being increasing in α when $\gamma < 1$ and independent of α when $\gamma = 1$. The first statement in part (ii) follows from (3.18), $F(w_t) = 0$ and $A(w_t)$ being decreasing in $u^\top \Sigma u$ when $\gamma > 1$. The second statement in part (ii) follows from the former two properties and $A(w_t)$ being increasing in $u^\top \Sigma u$ when $\gamma < 1$. The third statement in part (ii) follows from (3.19), $F(w_t) = 0$ and $A(w_t)$ being increasing in α when $\gamma < 1$. The result for long-lived hedgers follows from (3.18), $\lim_{w_t \to 0} F'(w_t) w_t = 0$, and because the asymptotic behavior of $A(w_t)$ for w_t close to zero is the same as for short-lived hedgers. ∎

Proof of Proposition 3.5: Substituting (3.17), (3.18) and (A.14) into (A.4) and (A.5), we can write the dynamics of arbitrageur wealth w_t as

$$dw_t = \mu_{wt} dt + \sigma_{wt}^\top dB_t, \tag{A.78}$$

where

$$\mu_{wt} = \left(r - q(w_t)^{-\frac{1}{\gamma}}\right) w_t + \frac{\alpha^2 A(w_t)}{[\alpha + A(w_t) - F'(w_t)]^2} u^\top \Sigma u, \tag{A.79}$$

$$\sigma_{wt} = \frac{\alpha}{\alpha + A(w_t) - F'(w_t)} \sigma u. \tag{A.80}$$

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

$$-[\mu_{wt} d(w_t)]' + \frac{1}{2}[\sigma_{wt}^\top \sigma_{wt} d(w_t)]'' = 0 \tag{A.81}$$

over $(0, \infty)$, with the boundary condition

$$\lim_{w_t \to \infty} \left[-\mu_{wt} d(w_t) + \frac{1}{2}[\sigma_{wt}^\top \sigma_{wt} d(w_t)]' \right] = 0 \tag{A.82}$$

(see Bogachev, Krylov, Röckner, and Shaposhnikov (2015)). Integrating (A.81) using (A.82) yields the ODE

$$-\mu_{wt} d(w_t) + \frac{1}{2}[\sigma_{wt}^\top \sigma_{wt} d(w_t)]' = 0. \tag{A.83}$$

Setting $D(w_t) \equiv \sigma_{wt}^\top \sigma_{wt} d(w_t)$, we can write (A.83) as

$$\frac{D'(w_t)}{D(w_t)} = \frac{2\mu_{wt}}{\sigma_{wt}^\top \sigma_{wt}}. \tag{A.84}$$

Integrating between one and w_t, we find

$$d(w_t) = D(1) \frac{\exp\left[\int_1^{w_t} \frac{2\mu_{wt}}{\sigma_{wt}^\top \sigma_{wt}} d\hat{w}_t\right]}{\sigma_{wt}^\top \sigma_{wt}}. \tag{A.85}$$

We can determine the multiplicative constant $D(1)$ by the requirement that $d(w_t)$ must integrate to one, i.e.,

$$\int_0^\infty d(w_t) dw_t = D(1) \int_0^\infty \frac{\exp\left[\int_1^{w_t} \frac{2\mu_{wt}}{\sigma_{wt}^\top \sigma_{wt}} d\hat{w}_t\right]}{\sigma_{wt}^\top \sigma_{wt}} dw_t = 1. \tag{A.86}$$

Equation (A.86) determines a positive $D(1)$, and hence a positive $d(w_t)$, if the integral multiplying $D(1)$ is finite. If the integral is infinite, then (A.86) implies that $D(1) = 0$, and the stationary distribution does not have a density but is concentrated at zero. The integral multiplying $D(1)$ is infinite when the integrand converges to infinity at a fast enough rate when w_t goes to zero, or does not converge to zero at a fast enough rate when w_t goes to infinity.

Substituting (μ_{wt}, σ_{wt}) from (A.79) and (A.80) into (A.85), we can write (A.85) as

$$
\begin{aligned}
d(w_t) &= D(1)\frac{\exp\left[\int_1^{w_t}\frac{2\left[\left(r-q(\hat{w}_t)^{-\frac{1}{\gamma}}\right)\hat{w}_t+\frac{\alpha^2 A(\hat{w}_t)}{[\alpha+A(\hat{w}_t)-F'(\hat{w}_t)]^2}u^\top\Sigma u\right]}{\frac{\alpha^2}{[\alpha+A(\hat{w}_t)-F'(\hat{w}_t)]^2}u^\top\Sigma u}d\hat{w}_t\right]}{\frac{\alpha^2}{[\alpha+A(w_t)-F'(w_t)]^2}u^\top\Sigma u} \\
&= D(1)\frac{\exp\left[\int_1^{w_t}\frac{2\left[-\frac{\rho-r}{A(\hat{w}_t)}\left(1+\frac{z\left(A'(\hat{w}_t)+A(\hat{w}_t)^2\right)}{[\alpha+A(\hat{w}_t)-F'(\hat{w}_t)]^2}\right)+\frac{\alpha^2 A(\hat{w}_t)}{[\alpha+A(\hat{w}_t)-F'(\hat{w}_t)]^2}u^\top\Sigma u\right]}{\frac{\alpha^2}{[\alpha+A(\hat{w}_t)-F'(\hat{w}_t)]^2}u^\top\Sigma u}d\hat{w}_t\right]}{\frac{\alpha^2}{[\alpha+A(w_t)-F'(w_t)]^2}u^\top\Sigma u} \\
&= D(1)\frac{\exp\left[\int_1^{w_t}\left(-\frac{A'(\hat{w}_t)}{A(\hat{w}_t)}+A(\hat{w}_t)-\frac{[\alpha+A(\hat{w}_t)-F'(\hat{w}_t)]^2}{zA(\hat{w}_t)}\right)d\hat{w}_t\right]}{\frac{\alpha^2}{[\alpha+A(w_t)-F'(w_t)]^2}u^\top\Sigma u} \\
&= \frac{D(1)A(1)}{\alpha^2 u^\top\Sigma u}\frac{[\alpha+A(w_t)-F'(w_t)]^2}{A(w_t)}\exp\left[\int_1^{w_t}\left(A(\hat{w}_t)-\frac{[\alpha+A(\hat{w}_t)-F'(\hat{w}_t)]^2}{zA(\hat{w}_t)}\right)d\hat{w}_t\right],
\end{aligned}
\tag{A.87}
$$

where the second step follows from (3.21), the third from (3.20), and the fourth by noting that

$$
\exp\left[\int_1^{w_t}\left(-\frac{A'(\hat{w}_t)}{A(\hat{w}_t)}\right)d\hat{w}_t\right] = \exp\{-\log[A(w_t)]+\log[A(1)]\} = \frac{A(1)}{A(w_t)}.
$$

Substituting (A.87) into (A.86) to solve for $D(1)$, and substituting back into (A.87), we find (3.24).

Since $\lim_{w_t\to 0} A(w_t)w_t = \max\{\gamma, K\}$ and $\lim_{w_t\to 0} F'(w_t)w_t = 0$, for w_t close to zero the numerator of (3.24) is bounded above and below by functions of the form

$$
\frac{\max\{\gamma,K\}}{w_t}\exp\left[\int_1^{w_t}\Psi_0\frac{z-1}{z}\frac{\max\{\gamma,K\}}{\hat{w}_t}d\hat{w}_t\right] = \frac{\max\{\gamma,K\}}{w_t}w_t^{\Psi_0\max\{\gamma,K\}\frac{z-1}{z}},
$$

where Ψ_0 is a positive constant. These functions are integrable at zero if $z>1$, and not integrable at zero if $z<1$. Hence, the numerator of (3.24) has the same properties. Since $\lim_{w_t\to\infty} A(w_t)w_t = \gamma$ and $\lim_{w_t\to\infty} F'(w_t)w_t = 0$, for w_t close to infinity the numerator of (3.24) is bounded above by a function of the form

$$
\frac{\alpha^2 w_t}{\gamma}\exp\left[\int_1^{w_t}-\Psi_\infty\frac{\alpha^2\hat{w}_t}{\gamma z}d\hat{w}_t\right] = \frac{\alpha^2 w_t}{\gamma}\exp\left[-\Psi_\infty\frac{\alpha^2(w_t^2-1)}{2\gamma z}\right],
$$

where Ψ_∞ is a positive constant. This function is integrable at infinity, and hence the numerator of (3.24) has the same property. Therefore, the integral in the denominator of (3.24) is finite if $z>1$ and infinite if $z<1$. ∎

Proof of Proposition 3.6: For $F(w_t) = 0$ and $A(w_t) = \frac{1}{w_t}$, (3.24) becomes

$$d(w_t) = \frac{(\alpha w_t + 1)^2 w_t^{-\frac{1}{z}} \exp\left(-\frac{1}{2z}\left(\alpha^2 w_t^2 + 4\alpha w_t\right)\right)}{\int_0^\infty (\alpha w_t + 1)^2 w_t^{-\frac{1}{z}} \exp\left(-\frac{1}{2z}\left(\alpha^2 w_t^2 + 4\alpha w_t\right)\right) dw_t}. \tag{A.88}$$

The derivative $d'(w_t)$ has the same sign as the derivative of the numerator. The latter derivative is

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

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

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

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

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

If $z < \frac{3}{2}$, then $\Phi'(x) > 0$ for all $x > 0$, and hence $\Phi(x) > 0$ for all $x > 0$. If $z > \frac{3}{2}$, then $\Phi'(x)$ 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 $\Phi(x_1') > 0$ then $\Phi(x) > 0$ for all $x > 0$, and if $\Phi(x_1') < 0$ then $\Phi(x)$ has two positive roots $x_1 < x_1' < x_2$ and is positive outside the roots and negative inside. Since

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

$\Phi(x_1')$ is positive if

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

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 $w_t \in (0, m_1) \cup (m_2, \infty)$ and positive for $w_t \in (m_1, m_2)$, where $m_i \equiv \frac{x_i}{\alpha}$ for $i = 1, 2$. This proves Parts (i) and (ii).

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

$$\frac{\partial^2 \log\left[d(w_t, \theta)\right]}{\partial\theta\partial w_t} > 0. \tag{A.89}$$

Using (A.88), we find

$$\frac{\partial \log[d(w_t)]}{\partial w_t} = \frac{2\alpha}{\alpha w_t + 1} - \frac{1}{z w_t} - \frac{1}{z}\left(\alpha^2 w_t + 2\alpha\right). \tag{A.90}$$

An increase in α (which also affects z from (3.20)) raises the right-hand side of (A.90). Therefore, $d(w_t)$ satisfies (A.89) with respect to α. An increase in z also raises the right-hand side of (A.90). Therefore, $d(w_t)$ satisfies (A.89) with respect to $u^\top \Sigma u$. This proves Part (iii). ∎

B Proofs of the Results in Section 4

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

$$dv_t = (rv_t - \bar{c}_t)dt + X_t^\top(\mu_{St} + \bar{D} - rS_t)dt + u^\top \bar{D} dt + \left(X_t^\top(\sigma_{St} + \sigma)^\top + u^\top \sigma^\top\right) dB_t, \tag{B.1}$$

$$dw_t = (rw_t - c_t)dt + Y_t^\top(\mu_{St} + \bar{D} - rS_t)dt + Y_t^\top(\sigma_{St} + \sigma)^\top dB_t, \tag{B.2}$$

respectively. If S_t, X_t, and Y_t satisfy (4.7), (4.8), and (4.9), then (B.1) is identical to (3.3), and (B.2) to (3.8). 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 (3.15), then X_t and Y_t satisfy the market-clearing equation (4.6) because of (4.8) and (4.9). Since (B.1) is identical to (3.3), and (B.2) to (3.8), 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 market prices η_t of the Brownian risks in the two equilibria are $(\sigma^\top)^{-1}(\bar{D} - \pi_t)$ and $\left((\sigma_{St} + \sigma)^\top\right)^{-1}(\mu_{St} + \bar{D} - rS_t)$, and are the same because of (4.7). The arbitrageurs' Sharpe ratios in the two equilibria are $\frac{y_t^\top(\bar{D} - \pi_t)}{\sqrt{y_t^\top \sigma^\top \sigma y_t}}$ and $\frac{Y_t^\top(\mu_{St} + \bar{D} - rS_t)}{\sqrt{Y_t^\top(\sigma_{St} + \sigma)^\top(\sigma_{St} + \sigma)Y_t}}$, and are the same because of (4.7) and (4.9). ∎

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

$$\begin{aligned}\mu_{St} &= \mu_{wt} S'(w_t) + \frac{1}{2}\sigma_{wt}^\top \sigma_{wt} S''(w_t) \\ &= \left(r - q(w_t)^{-\frac{1}{\gamma}}\right) S'(w_t) w_t + \frac{\alpha^2 u^\top \Sigma u}{[\alpha + A(w_t) - F'(w_t)]^2}\left[A(w_t)S'(w_t) + \frac{1}{2}S''(w_t)\right], \end{aligned} \tag{B.3}$$

where the second step follows from (A.79) and (A.80), and

$$\begin{aligned}\sigma_{St} &= \sigma_{wt} S'(w_t)^\top \\ &= \frac{\alpha}{\alpha + A(w_t) - F'(w_t)} \sigma u S'(w_t)^\top, \end{aligned} \tag{B.4}$$

where the second step follows from (A.80). Multiplying (4.7) from the left by $(\sigma_{St} + \sigma)^\top$, and using (3.17), we find

$$\mu_{St} + \bar{D} - rS_t = \frac{\alpha A(w_t)}{\alpha + A(w_t) - F'(w_t)} (\sigma_{St} + \sigma)^\top \sigma u. \tag{B.5}$$

Substituting (μ_{St}, σ_{St}) from (B.3) and (B.4) into (B.5), we find the ODE

$$\left(r - q(w_t)^{-\frac{1}{\gamma}}\right) S'(w_t) w_t + \frac{\alpha^2 u^\top \Sigma u}{2[\alpha + A(w_t) - F'(w_t)]^2} S''(w_t) + \bar{D} - rS(w_t) = \frac{\alpha A(w_t)}{\alpha + A(w_t) - F'(w_t)} \Sigma u. \tag{B.6}$$

Setting $S(w_t) \equiv \frac{\bar{D}}{r} + \hat{S}(w_t)$, we find that the resulting ODE equates linear terms in $\hat{S}(w_t)$, $\hat{S}'(w_t)$ and $\hat{S}''(w_t)$ to a scalar times Σu. Hence, $\hat{S}(w_t)$ must be collinear to Σu, which means that $S(w_t)$ must have the form in (4.10). Substituting (4.10) into (B.6), we find that $g(w_t)$ solves the ODE (4.11).

Substituting μ_{St} from (B.3) into (4.2), and using (4.10) and (4.11), we can write expected excess returns as (4.12). Substituting σ_{St} from (B.4) into (4.3), and using (4.10), we can write the covariance matrix of returns as (4.13). ∎

Proof of Theorem 4.1: We start with the case where hedgers are short-lived. Since $q(w_t)$ and $A(w_t)$ are continuous and positive in $(0, \infty)$, the functions multiplying $g(w_t)$, $g'(w_t)$ and $g''(w_t)$ in (4.11) are continuous. Since, in addition, (4.11) is a linear ODE, it has a unique solution over any interval $[\epsilon, M] \subset (0, \infty)$ with initial conditions $g(\epsilon)$ and $g'(\epsilon)$ (Murray and Miller (2013)). That solution can be extended over $(0, \infty)$. We next derive properties of solutions to (4.11), as well as the existence result, through a number of lemmas that parallel those in the proof of Theorem 3.1.

Lemma B.1 (Limits at zero and infinity) *Consider a solution $g(w_t)$ to (4.11), defined over the interval $(0,\infty)$. If the limits of $g(w_t)$ at zero and infinity are finite, then they are equal to $-\frac{\alpha}{r}$ and zero, respectively.*

Proof: We first derive the limit at zero. Suppose that $\lim_{w_t\to 0} g'(w_t)w_t$ exists, in which case it is zero (as shown in the proof of Theorem 3.1 in the case of $R(w_t)$). Since $\lim_{w_t\to 0} g(w_t)$ is assumed to exist, (4.11) implies that $\lim_{w_t\to 0} \frac{g''(w_t)}{[\alpha+A(w_t)]^2}$ exists. Since $\lim_{w_t\to 0} A(w_t)w_t = \max\{\gamma, K\}$,

$$\lim_{w_t\to 0} \frac{g''(w_t)}{[\alpha + A(w_t)]^2} = \frac{\lim_{w_t\to 0} g''(w_t)w_t^2}{(\max\{\gamma,K\})^2},$$

and hence $\lim_{w_t\to 0} g''(w_t)w_t^2$ exists. As shown in the proof of Theorem 3.1 (in the case of $F(w_t)$), if $\lim_{w_t\to 0} g''(w_t)w_t^2$ exists, it is zero. Taking the limit of both sides of (4.11) when w_t goes to zero and using $\lim_{w_t\to 0} g'(w_t)w_t = \lim_{w_t\to 0} g''(w_t)w_t^2 = 0$, $\lim_{w_t\to 0} A(w_t)w_t = \max\{\gamma, K\}$ and $\lim_{w_t\to 0} q(w_t)^{-\frac{1}{\gamma}} \in (0,\infty)$, we find $\lim_{w_t\to 0} g(w_t) = -\frac{\alpha}{r}$.

To complete the proof for the limit at zero, we need to show that $\lim_{w_t\to 0} g'(w_t)w_t$ exists. We proceed by contradiction and assume that $\lim_{w_t\to 0} g'(w_t)w_t$ does not exist, and hence $\limsup_{w_t\to 0} g'(w_t)w_t > \liminf_{w_t\to 0} g'(w_t)w_t$. Since $g'(w_t)w_t$ oscillates between values close to $\limsup_{w_t\to 0} g'(w_t)w_t$ and values close to $\liminf_{w_t\to 0} g'(w_t)w_t$, there exists $\xi \in (\liminf_{w_t\to 0} g'(w_t)w_t, \limsup_{w_t\to 0} g'(w_t)w_t)$ and a sequence $\{w_{tn}\}_{n\in\mathbb{N}}$ converging to zero such that $g'(w_{tn})w_{tn} = \xi$ and $[g'(w_t)w_t]'\big|_{w_t=w_{tn}}$ alternates between being non-positive and non-negative. Since

$$\left[g'(w_t)w_t\right]'\Big|_{w_t=w_{tn}} = g''(w_{tn})w_{tn} + g'(w_{tn}) = \frac{g''(w_{tn})w_{tn}^2 + \xi}{w_{tn}},$$

where the second step follows from $g'(w_{tn})w_{tn} = \xi$,

$$g''(w_{tn})w_{tn}^2 + \xi$$

must also alternate between being non-positive and non-negative. Taking the limit of (4.11) along the sequence $\{w_{tn}\}_{n\in\mathbb{N}}$, however, we find

$$\frac{\alpha^2 u^\top \Sigma u}{2(\max\{\gamma,K\})^2} \lim_{w_t\to 0} g''(w_{tn})w_{tn}^2 + \left(r - \lim_{w_t\to 0} q(w_t)^{-\frac{1}{\gamma}}\right)\xi - r\lim_{w_t\to 0} g(w_t) = \alpha.$$

Hence, $g''(w_{tn})w_{tn}^2$ converges to a finite limit, which means that ξ can be chosen so that $g''(w_{tn})w_{tn}^2 + \xi$ can be non-zero and with a sign that does not change, a contradiction.

We next derive the limit at infinity. Suppose that $\lim_{w_t\to\infty} g'(w_t)w_t$ exists, in which case it is zero. The same argument as for the limit at zero implies that $\lim_{w_t\to\infty} g''(w_t)$ exists. That limit has to be zero since $\lim_{w_t\to\infty} g'(w_t)w_t = 0$. Taking the limit of both sides of (4.11) when w_t goes to zero and using $\lim_{w_t\to\infty} g'(w_t)w_t = \lim_{w_t\to\infty} g''(w_t) = 0$, $\lim_{w_t\to\infty} A(w_t) = 0$ and $\lim_{w_t\to\infty} q(w_t)^{-\frac{1}{\gamma}} \in (0,\infty)$, we find $\lim_{w_t\to\infty} g(w_t) = 0$. The existence of $\lim_{w_t\to\infty} g'(w_t)w_t$ follows by adapting the contradiction argument used for the limit at zero. ∎

Lemma B.2 (Single crossing of solutions) *Consider two solutions $g_1(w_t)$ and $g_2(w_t)$ to (4.11) with initial conditions $g_1(\epsilon) = g_2(\epsilon)$ and $g_1'(\epsilon) > g_2'(\epsilon)$ for $\epsilon > 0$. The solutions compare as follows:*

- *$g_1(w_t) > g_2(w_t)$ for all $w_t \in (\epsilon,\infty)$, and $g_1(w_t) < g_2(w_t)$ for all $w_t \in (0,\epsilon)$.*
- *$g_1'(w_t) > g_2'(w_t)$ for all $w_t \in (0,\infty)$.*

Proof: We first show the inequalities for $w_t > \epsilon$. Since $g_1(\epsilon) = g_2(\epsilon)$ and $g_1'(\epsilon) > g_2'(\epsilon)$, $g_1(w_t) > g_2(w_t)$ for w_t close to and larger than ϵ. Proceeding by contradiction, suppose that there exists $w_t > \epsilon$ such that $g_1(w_t) \le g_2(w_t)$ or $g_1'(w_t) \le g_2'(w_t)$. The infimum m within that set is strictly larger than ϵ since $g_1(w_t) > g_2(w_t)$ and $g_1'(w_t) > g_2'(w_t)$ for w_t close to and larger than ϵ. Since $g_1(\epsilon) = g_2(\epsilon)$ and $g_1'(w_t) > g_2'(w_t)$ for all $w_t \in (\epsilon, m)$, $g_1(m) > g_2(m)$. Hence, $g_1'(m)$ must be equal to $g_2'(m)$. Since $g_1'(w_t) > g_2'(w_t)$ for all $w_t \in (\epsilon, m)$, $g_1''(m) \le g_2''(m)$. Equation (4.11) then implies $g_1(m) \le g_2(m)$, a contradiction. Therefore, $g_1(w_t) > g_2(w_t)$ and $g_1'(w_t) > g_2'(w_t)$ for all $w_t \in (\epsilon,\infty)$. The inequalities for $w_t < \epsilon$ follow from a similar argument (developed in the case of $Q(w_t)$ and $R(w_t)$ in Lemma A.2). ∎

Lemma B.3 (Boundary conditions over finite interval) *For any $\epsilon > 0$ and $M > \epsilon$, there exists a unique solution to (4.11) defined over the interval $(0,\infty)$ that satisfies $g(\epsilon) = -\frac{\alpha}{r}$ and $g(M) = 0$.*

Proof: We will consider solutions to (4.11) with $g(\epsilon) = -\frac{\alpha}{r}$, and show that there exists a unique $g'(\epsilon)$ such that $g(M) = 0$. For any $\ell \in (-\infty,\infty)$, the linear ODE (4.11) has a unique solution over $[\epsilon, M]$ with initial conditions $g(\epsilon) = -\frac{\alpha}{r}$ and $g'(\epsilon) = \ell$, and that solution can be extended over $(0,\infty)$. We denote the solution derived for $g'(\epsilon) = \ell$ by $g_\ell(w_t)$.

Consider the solutions $g_0(w_t)$ and $g_1(w_t)$. Since the ODE (4.11) is linear, the function $h(w_t) \equiv g_0(w_t) + \ell\,[g_1(w_t) - g_0(w_t)]$ is also a solution. Since, in addition, $h(\epsilon) = -\frac{\alpha}{r}$ and $h'(\epsilon) = \ell$, $h(w_t)$ coincides with the solution $g_\ell(w_t)$.

Since $g_0(\epsilon) = g_1(\epsilon) = -\frac{\alpha}{r}$ and $g_1'(\epsilon) > g_0'(\epsilon)$, Lemma B.2 implies $g_1(w_t) - g_0(w_t) > 0$ for all $w_t \in (\epsilon, \infty)$. Therefore, the function $\ell \longrightarrow g_0(M) + \ell\,[g_1(M) - g_0(M)] = g_\ell(M)$ is invertible, and so there exists a unique ℓ such that $g_\ell(M) = 0$. The function $g_\ell(w_t)$ corresponding to that ℓ is the required solution. ∎

Lemma B.4 (Negative and increasing $g(w_t)$) *For any $\epsilon > 0$ and $M > \epsilon$, the solution to (4.11) constructed in Lemma B.3 is negative for all w_t in $[\epsilon, M)$ and increasing.*

Proof: We first show that the ℓ determined in Lemma B.3 is positive. Since $g_\ell(M) = 0$, $g_1(M) - g_0(M) > 0$ and $g_\ell(M) = g_0(M) + \ell\,[g_1(M) - g_0(M)]$, we need to show that $g_0(M) < 0$. We will show the stronger result that $g_0(M) < -\frac{\alpha}{r}$. Setting $w_t = \epsilon$, $g_0(\epsilon) = -\frac{\alpha}{r}$ and $g_0'(\epsilon) = 0$ in (4.11), we find $g_0''(\epsilon) < 0$. Hence, $g_0'(w_t) < 0$ and $g_0(w_t) < -\frac{\alpha}{r}$ for w_t close to and larger than ϵ. Proceeding by contradiction, suppose that there exists $w_t > \epsilon$ such that $g_0(w_t) \geq -\frac{\alpha}{r}$ or $g_0'(w_t) \geq 0$. The infimum m within that set is strictly larger than ϵ since $g_0(w_t) < -\frac{\alpha}{r}$ and $g_0'(w_t) < 0$ for w_t close to and larger than ϵ. Since $g_0(\epsilon) = -\frac{\alpha}{r}$ and $g_0'(w_t) < 0$ for all $w_t \in (\epsilon, m)$, $g_0(m) < -\frac{\alpha}{r}$. Hence, $g_0'(m)$ must be equal to zero. Since $g_0'(w_t) < 0$ for all $w_t \in (\epsilon, m)$, $g_0''(m) \geq 0$. Since, however, $g_0(m) < -\frac{\alpha}{r}$ and $g'(m) = 0$, (4.11) implies $g_0''(m) < 0$, a contradiction. Therefore, $g_0(w_t) < -\frac{\alpha}{r}$ and $g_0'(w_t) < 0$ for all $w_t \in (\epsilon, \infty)$. This in turn implies $g_0(M) < 0$ and $\ell > 0$.

We next show that $g_\ell(w_t) < g_\ell(\hat{w}_t)$ for all $w_t < \hat{w}_t$ in $[\epsilon, M]$. Suppose, by contradiction, that there exist $m_2 > m_1$ in $[\epsilon, M]$ such that $g_\ell(m_1) \geq g_\ell(m_2)$. If $g_\ell(m_1) > g_\ell(m_2)$, then there exists $m_3 \in (m_1, m_2)$ such that $g_\ell(m_3) = \frac{g_\ell(m_1)+g_\ell(m_2)}{2}$. Since $g_\ell'(\epsilon) = \ell > 0$ and $g_\ell(m_1) > g_\ell(m_3)$, $g_\ell(w_t)$ reaches its maximum value over $[\epsilon, m_3]$ at an interior point m. Setting $g_\ell'(m) = 0$ in (4.11), we find that $g_\ell''(m)$ has the same sign as $\frac{A(m)}{\alpha + A(m)} + r g_\ell(m)$. Since $g_\ell''(m) \leq 0$ at a maximum,

$$\frac{A(m)}{\alpha + A(m)} + r g_\ell(m) \leq 0 \Rightarrow g_\ell(m) \leq -\frac{A(m)}{r[\alpha + A(m)]} < 0. \tag{B.7}$$

Since $g_\ell(m_3) < g_\ell(m) < 0 = g_\ell(M)$ and $g_\ell(m_3) > g_\ell(m_2)$, $g_\ell(w_t)$ reaches its minimum value in $[m_3, M]$ at an interior point $\hat{m} > m_3$. Setting $g_\ell'(\hat{m}) = 0$ in (4.11), we find that $g_\ell''(\hat{m})$ has the same sign as $\frac{A(\hat{m})}{\alpha+A(\hat{m})}+rg_\ell(\hat{m})$. Since $g''_\ell(\hat{m}) \geq 0$ at a minimum,

$$\frac{A(\hat{m})}{\alpha+A(\hat{m})}+rg_\ell(\hat{m}) \geq 0 \Rightarrow g_\ell(\hat{m}) \geq -\frac{A(\hat{m})}{r[\alpha+A(\hat{m})]}. \tag{B.8}$$

Since $A(w_t)$ is decreasing, (B.7) and (B.8) imply $g_\ell(m) < g_\ell(\hat{m})$. Since, however, m is the maximizer of $g_\ell(w_t)$ over $[\epsilon, m_3]$ and $\hat{m}$ is the minimizer of $g_\ell(w_t)$ over $[m_3, M]$, $g_\ell(m) > g_\ell(\hat{m})$, a contradiction. If $g_\ell(m_1) = g_\ell(m_2)$, a contradiction can be derived through the same argument, by distinguishing the case where there exists $w_t \in (m_1, m_2)$ such that $g_\ell(w_t) \neq g_\ell(m_1) = g_\ell(m_2)$ and replacing m_1 or m_2 by w_t, and the case where $g_\ell(w_t) = g_\ell(m_1) = g_\ell(m_2)$ for all $w_t \in (m_1, m_2)$ and taking $\frac{m_1+m_2}{2}$ for m_3. Therefore, $g_\ell(w_t) < g_\ell(\hat{w}_t)$ for all $w_t < \hat{w}_t$ in $[\epsilon, M]$. Since $g_\ell(M) = 0$, $g_\ell(w_t) < 0$ for all $w_t \in [\epsilon, M)$. ∎

Lemma B.5 (Boundary conditions over $[0,\infty]$**)** *A solution to (4.11), defined over the interval* $(0,\infty)$*, and with finite limits at zero and infinity, exists. That solution is negative and increasing.*

Proof: We will construct the solution as the simple limit of solutions with boundary conditions at ϵ and M. Denote the solution constructed in Lemma B.3 by $g_{\epsilon,M}(w_t)$. Since for $M_2 > M_1$, $g_{\epsilon,M_2}(M_1) < 0 = g_{\epsilon,M_1}(M_1)$ (as implied by Lemma B.4), Lemma B.2 implies $g'_{\epsilon,M_1}(\epsilon) > g'_{\epsilon,M_2}(\epsilon)$, and hence $g_{\epsilon,M_1}(w_t) > g_{\epsilon,M_2}(w_t)$ for all $w_t \in (\epsilon,\infty)$. This means that the function $M \to g_{\epsilon,M}(w_t)$, defined for given $w_t > \epsilon$ and for $M > w_t$, is decreasing. Since that function is bounded below by $-\frac{\alpha}{r}$ (as implied by Lemma B.4), it converges to a finite limit, denoted by $g_\epsilon(w_t)$, when M goes to infinity. That limit satisfies $0 \geq g_\epsilon(w_t) \geq -\frac{\alpha}{r}$ for all $w_t \in (\epsilon,\infty)$ and $g_\epsilon(w_t) \leq g_\epsilon(\hat{w}_t)$ for all $w_t < \hat{w}_t$ in (ϵ,∞) because these inequalities hold (strictly) for $g_{\epsilon,M}(w_t)$ over (ϵ, M) (Lemma B.4).

We next take the limit of $g_\epsilon(w_t)$ when ϵ goes to zero. Since for $\epsilon_2 < \epsilon_1$ and for all $M > \epsilon_1$, $g_{\epsilon_2,M}(\epsilon_1) > -\frac{\alpha}{r} = g_{\epsilon_1,M}(\epsilon_1)$ (as implied by Lemma B.4), Lemma B.2 (applied with M rather than ϵ as starting value) implies $g'_{\epsilon_2,M}(M) < g'_{\epsilon_1,M}(M)$, and hence $g_{\epsilon_2,M}(w_t) > g_{\epsilon_1,M}(w_t)$ for all $w_t \in (0, M)$. Taking limits when M goes to infinity, we find $g_{\epsilon_2}(w_t) \geq g_{\epsilon_1}(w_t)$ for all $w_t \in (\epsilon_1,\infty)$. Hence, the function $\epsilon \to g_\epsilon(w_t)$, defined for given $w_t > \epsilon$, is increasing. Since that function is bounded above by zero, it converges to a finite limit, denoted by $g(w_t)$, when ϵ goes to zero. That limit satisfies $0 \geq g(w_t) \geq -\frac{\alpha}{r}$ for all $w_t \in (0,\infty)$ and $g(w_t) \leq g(\hat{w}_t)$ for all $w_t < \hat{w}_t$ in $(0,\infty)$ because these inequalities hold for $g_\epsilon(w_t)$ over (ϵ,∞).

Following a similar argument as in Lemma A.5, we can show that $g(w_t)$, viewed as a function of w_t, solves the ODE (4.11). Since $0 \geq g(w_t) \geq -\frac{\alpha}{r}$ for all $w_t \in (0,\infty)$ and $g(w_t) \leq g(\hat{w}_t)$ for all $w_t < \hat{w}_t$ in $(0,\infty)$, $g(w_t)$ has finite limits at zero and infinity. To show that $g(w_t)$ is negative and increasing, we need to show that the inequalities $0 \geq g(w_t) \geq -\frac{\alpha}{r}$ for all $w_t \in (0,\infty)$ and $g(w_t) \leq g(\hat{w}_t)$ for all $w_t < \hat{w}_t$ in $(0,\infty)$ are strict. This can be done following a similar argument as in Lemma B.4. ∎

When hedgers are long-lived, the analysis in Lemma B.1 carries through, by observing additionally that $\lim_{w_t\to 0} F'(w_t)w_t = \lim_{w_t\to\infty} F'(w_t)w_t = 0$. ∎

Proof of Proposition 4.2: Showing that the effects of w_t on variance, covariance and correlation converge to zero when w_t goes to zero and to infinity amounts to showing that $\lim_{w_t\to 0} f(w_t) = \lim_{w_t\to\infty} f(w_t) = 0$. To show that $\lim_{w_t\to 0} f(w_t) = 0$, we multiply the numerator and denominator of $f(w_t)$ by w_t:

$$\lim_{w_t\to 0} f(w_t) = \frac{\alpha \lim_{w_t\to 0} g'(w_t)w_t}{\alpha \lim_{w_t\to 0} w_t + \lim_{w_t\to 0} A(w_t)w_t - \lim_{w_t\to 0} F'(w_t)w_t}. \tag{B.9}$$

Since $\lim_{w_t\to 0} A(w_t)w_t = \max\{\gamma, K\}$, $\lim_{w_t\to 0} F'(w_t)w_t = 0$ and $\lim_{w_t\to 0} g'(w_t)w_t = 0$, (B.9) implies $\lim_{w_t\to 0} f(w_t) = 0$. To show that $\lim_{w_t\to\infty} f(w_t) = 0$, we follow the same procedure and note that $\lim_{w_t\to\infty} A(w_t)w_t = \gamma$, $\lim_{w_t\to\infty} F'(w_t)w_t = 0$ and $\lim_{w_t\to\infty} g'(w_t)w_t = 0$.

We next show that when hedgers are short-lived and arbitrageurs have logarithmic preferences, $f(w_t)$ is hump-shaped. Since $\lim_{w_t\to 0} f(w_t) = \lim_{w_t\to\infty} f(w_t) = 0$, $f(w_t) \geq 0$ for all $w_t \in (0,\infty)$ and $f(w_t) > 0$ for at least some $w_t \in (0,\infty)$, there exists $w_t \in (0,\infty)$ at which $f(w_t)$ is maximized. If $f(w_t)$ is not hump-shaped, then we can proceed as in Lemma A.7 and choose $\xi > 0$ and $(\hat{m}_i, \hat{\hat{m}}_i)$ for $i = 1,2,3$, such that $\hat{m}_1 < \hat{\hat{m}}_1 \leq \hat{m}_2 < \hat{\hat{m}}_2 \leq \hat{m}_3 < \hat{\hat{m}}_3$, $f(\hat{m}_i) = f(\hat{\hat{m}}_i) = \xi$ for $i = 1,2,3$, $f'(\hat{m}_i) > 0$ and $f'(\hat{\hat{m}}_i) < 0$ for $i = 1,3$, $f'(\hat{m}_2) < 0$ and $f'(\hat{\hat{m}}_2) > 0$, $f(w_t) > \xi$ for all $w_t \in (\hat{m}_i, \hat{\hat{m}}_i)$ for $i = 1,3$, and $f(w_t) < \xi$ for all $w_t \in (\hat{m}_2, \hat{\hat{m}}_2)$.

For any γ, the derivative of $f(w_t)$ has the same sign as

$$\begin{aligned}
&g''(w_t)[\alpha + A(w_t)] - g'(w_t)A'(w_t) \\
&= \frac{[\alpha + A(w_t)]^3}{\alpha^2 u^\top \Sigma u}\left[\frac{\alpha A(w_t)}{\alpha + A(w_t)} + rg(w_t) + \left(q(w_t)^{-\frac{1}{\gamma}} - r\right) g'(w_t)w_t\right] - g'(w_t)A'(w_t),
\end{aligned} \tag{B.10}$$

where the second step follows by substituting $g''(w_t)$ from (4.11). Since for $\gamma = 1$, $A(w_t) = \frac{1}{w_t}$ and $q(w_t) = \frac{1}{\rho}$, we can write (B.10) as

$$\frac{(\alpha w_t + 1)^3}{\alpha^2 u^\top \Sigma u w_t^3} \left[\frac{\alpha}{\alpha w_t + 1} + r g(w_t) + (\rho - r) g'(w_t) w_t \right] + \frac{g'(w_t)}{w_t^2}. \tag{B.11}$$

Using (B.11) and $f(\hat{m}_i) = f(\hat{\hat{m}}_i) = \xi$ for $i = 1, 2, 3$, we find $\hat{f}_\xi(\hat{m}_i) > 0$ and $\hat{f}_\xi(\hat{\hat{m}}_i) < 0$ for $i = 1, 3$, and $\hat{f}_\xi(\hat{m}_2) < 0$ and $\hat{f}_\xi(\hat{\hat{m}}_2) > 0$, where the function $\hat{f}_\xi(w_t)$ is defined by

$$\hat{f}_\xi(w_t) \equiv \frac{1}{\alpha u^\top \Sigma u} \left[\frac{\alpha}{\alpha w_t + 1} + r g(w_t) + \frac{(\rho - r)(\alpha w_t + 1)\xi}{\alpha} \right] + \frac{\xi}{(\alpha w_t + 1)^2}. \tag{B.12}$$

Consider next the function

$$\begin{aligned} \bar{\hat{f}}_\xi(w_t) &\equiv \hat{f}_\xi(w_t) - \frac{r}{\alpha u^\top \Sigma u} \left[g(w_t) - \frac{\xi}{\alpha} \left[\alpha w_t + \log(w_t) \right] \right] \\ &= \frac{1}{\alpha u^\top \Sigma u} \left[\frac{\alpha}{\alpha w_t + 1} + \frac{r\xi}{\alpha} \left[\alpha w_t + \log(w_t) \right] + \frac{(\rho - r)(\alpha w_t + 1)\xi}{\alpha} \right] + \frac{\xi}{(\alpha w_t + 1)^2}. \end{aligned} \tag{B.13}$$

Since $f(w_t) > \xi$ for all $w_t \in (\hat{m}_1, \hat{m}_1)$,

$$g'(w_t) > \frac{\xi}{\alpha} \left(\alpha + \frac{1}{w_t} \right) \Rightarrow g(\hat{\hat{m}}_1) - \frac{\xi}{\alpha} \left[\alpha \hat{\hat{m}}_1 + \log(\hat{\hat{m}}_1) \right] > g(\hat{m}_1) - \frac{\xi}{\alpha} \left[\alpha \hat{m}_1 + \log(\hat{m}_1) \right]. \tag{B.14}$$

Combining (B.14) with (B.13), $\hat{f}_\xi(\hat{m}_1) > 0$ and $\hat{f}_\xi(\hat{\hat{m}}_1) < 0$, we find $\bar{\hat{f}}_\xi(\hat{m}_1) > \bar{\hat{f}}_\xi(\hat{\hat{m}}_1)$. We likewise find $\bar{\hat{f}}_\xi(\hat{m}_2) < \bar{\hat{f}}_\xi(\hat{\hat{m}}_2)$ and $\bar{\hat{f}}_\xi(\hat{m}_3) > \bar{\hat{f}}_\xi(\hat{\hat{m}}_3)$. Hence, $\bar{\hat{f}}_\xi(w_t)$ is decreasing, increasing, and then decreasing. Since $\bar{\hat{f}}_\xi(w_t)$ converges to minus infinity when w_t goes to zero, it starts being increasing, and since it converges to infinity when w_t goes to infinity, it ends being increasing. Hence, its first derivative changes sign at least four times, from positive, to negative, to positive, to negative, to positive. Moreover, its second derivative changes sign at least three times, from negative, to positive, to negative, to positive. The first and second derivatives of $\bar{\hat{f}}_\xi(w_t)$ are

$$\bar{\hat{f}}'_\xi(w_t) = \frac{1}{\alpha u^\top \Sigma u} \left[-\frac{\alpha^2}{(\alpha w_t + 1)^2} + r\xi \left(1 + \frac{1}{\alpha w_t} \right) + (\rho - r)\xi \right] - \frac{2\alpha\xi}{(\alpha w_t + 1)^3}, \tag{B.15}$$

$$\bar{\hat{f}}''_\xi(w_t) = \frac{1}{\alpha u^\top \Sigma u} \left[\frac{2\alpha^2}{(\alpha w_t + 1)^3} - \frac{r\xi}{\alpha w_t^2} \right] + \frac{6\alpha^2 \xi}{(\alpha w_t + 1)^4}, \tag{B.16}$$

respectively. The second derivative has the same sign as the function $P(w_t) - \frac{r\xi}{\alpha}$, where

$$P(w_t) \equiv \frac{2\alpha^2 w_t^2}{(\alpha w_t + 1)^3} + \frac{6\alpha^3 u^\top \Sigma u \xi w_t^2}{(\alpha w_t + 1)^4}.$$

The derivative of $P(w_t)$ has the same sign as the expression

$$(2 - \alpha w_t)(1 + \alpha w_t) + 6\alpha u^\top \Sigma u \xi (1 - 2\alpha w_t),$$

which is quadratic in w_t, positive for $w_t = 0$, and has a unique positive root. Hence, $P(w_t)$ is hump-shaped, which means that $\bar{\hat{f}}''_\xi(w_t)$ can change sign at most twice, from negative, to positive, to negative. This yields a contradiction, and hence $f(w_t)$ is hump-shaped.

Parts (i) and (ii) of the proposition follow from (4.13) and $f(w_t)$ being hump-shaped. To show Part (iii), we use (4.13) to write the correlation as

$$\mathrm{Corr}_t(dR_{nt}, dR_{n't}) = \frac{f(w_t)\left[u^\top \Sigma u f(w_t) + 2\right](\Sigma u)_n (\Sigma u)_{n'} + \Sigma_{nn'}}{\sqrt{\left\{f(w_t)\left[u^\top \Sigma u f(w_t) + 2\right](\Sigma u)_n^2 + \Sigma_{nn}\right\}\left\{f(w_t)\left[u^\top \Sigma u f(w_t) + 2\right](\Sigma u)_{n'}^2 + \Sigma_{n'n'}\right\}}}. \tag{B.17}$$

Differentiating (B.17) with respect to $f(w_t)$, we find that $\mathrm{Corr}_t(dR_{nt}, dR_{n't})$ is increasing in $f(w_t)$ if (4.15) holds and is decreasing in $f(w_t)$ if (4.15) holds in the opposite direction. Part (iii) then follows from $f(w_t)$ being hump-shaped.

Proof of Proposition 4.3: We start by proving a lemma on the limit of $g'(w_t)$ when w_t goes to zero.

Lemma B.6 (Limit of $g'(w_t)$ at zero) *For the solution to (4.11) constructed in Lemma B.5, the limit of $g'(w_t)$ when w_t goes to zero is infinite if $\gamma < K$ and equal to* $\frac{\alpha^2\gamma}{(\rho - r)z\left[\gamma + \frac{\gamma}{z}\left(\frac{r}{\rho - r} + 1\right) - 1\right]}$ *if $\gamma > K$.*

Proof: We start with the case where hedgers are short-lived. We use the equation

$$\frac{\alpha^2 u^\top \Sigma u}{2[\alpha + A(w_t)]^2 w_t} g''(w_t) + \left(r - q(w_t)^{-\frac{1}{\gamma}}\right) g'(w_t) - r\frac{g(w_t) + \frac{\alpha}{r}}{w_t} = -\frac{\alpha^2}{[\alpha + A(w_t)]w_t}, \tag{B.18}$$

which can be derived from (4.11) by subtracting α from both sides and dividing by w_t. Consider first the case $\gamma > K$, and suppose that $g'(w_t)$ converges to a finite limit when w_t goes to zero.

Since $\lim_{w_t \to 0} g'(w_t)$ is finite and $\lim_{w_t \to 0} g(w_t) = -\frac{\alpha}{r}$, $\lim_{w_t \to 0} \frac{g(w_t)+\frac{\alpha}{r}}{w_t} = \lim_{w_t \to 0} g'(w_t)$. Since, $\lim_{w_t \to 0} A(w_t)w_t = \gamma$, $\lim_{w_t \to 0} \frac{\alpha^2}{[\alpha + A(w_t)]w_t} = \frac{\alpha^2}{\gamma}$. Since, in addition, $q(w_t)^{-\frac{1}{\gamma}}$ converges to a finite limit when w_t goes to zero, (B.18) implies that $g''(w_t)w_t$ must converge to a finite limit, which has to be zero because $\lim_{w_t \to 0} g'(w_t)$ is finite. Taking the limit of both sides of (B.18) when w_t goes to zero thus yields

$$\lim_{w_t \to 0} q(w_t)^{-\frac{1}{\gamma}} \lim_{w_t \to 0} g'(w_t) = \frac{\alpha^2}{\gamma}. \tag{B.19}$$

Substituting $\lim_{w_t \to 0} q(w_t)^{-\frac{1}{\gamma}} = \lim_{w_t \to 0} Q(w_t)$ from (A.28) into (B.19), we find that $\lim_{w_t \to 0} g'(w_t)$ is as in the lemma.

To complete the proof of the lemma for $\gamma > K$, we need to show that $g'(w_t)$ converges to a finite limit when w_t goes to zero, and does not have no limit or converge to infinity. Suppose, by contradiction, that $\lim_{w_t \to 0} g'(w_t)$ does not exist, and hence $\limsup_{w_t \to 0} g'(w_t)w_t > \liminf_{w_t \to 0} g'(w_t)w_t$. Since $g'(w_t)$ oscillates between values close to $\limsup_{w_t \to 0} g'(w_t)$ and values close to $\liminf_{w_t \to 0} g'(w_t)$, there exist $\xi > \hat{\xi}$ in $(\liminf_{w_t \to 0} g'(w_t), \limsup_{w_t \to 0} g'(w_t))$ and sequences $\{w_{tn}\}_{n \in \mathbb{N}}$ and $\{\hat{w}_{tn}\}_{n \in \mathbb{N}}$ converging to zero with the following properties. For the former sequence, $g'(w_{tn}) = \xi$, $g''(\hat{w}_{tn})$ alternates between being non-positive and non-negative, and $g'(w_t) \geq \xi$ for each interval in which $g''(\hat{w}_{tn})$ is non-negative at the lower end and non-positive at the upper end. For the latter sequence, $g'(\hat{w}_{tn}) = \hat{\xi}$, $g''(w_{tn})$ alternates between being non-positive and non-negative, and $g'(w_t) \leq \hat{\xi}$ for each interval in which $g''(w_{tn})$ is non-positive at the lower end and non-negative at the upper end. Denote by w_{tn} the element of the sequence corresponding to the upper end of the interval, in which case the element corresponding to the lower end is $w_{t,n+1}$. Using (4.11) and $g'(w_{tn}) = g'(w_{t,n+1}) = \xi$, we find $\hat{g}_\xi(w_{tn}) \leq 0$ and $\hat{g}_\xi(w_{t,n+1}) \geq 0$, where the function $\hat{g}_\xi(w_t)$ is defined by

$$\hat{g}_\xi(w_t) \equiv \frac{\alpha A(w_t)}{\alpha + A(w_t)} + rg(w_t) + \left(q(w_t)^{-\frac{1}{\gamma}} - r\right) \xi w_t. \tag{B.20}$$

Consider next the function

$$\bar{\hat{g}}_\xi(w_t) \equiv \hat{g}_\xi(w_t) - r[g(w_t) - \xi w_t] = \frac{\alpha A(w_t)}{\alpha + A(w_t)} + \xi q(w_t)^{-\frac{1}{\gamma}} w_t. \tag{B.21}$$

Since $g'(w_t) \geq \xi$ over the interval $[w_{t,n+1}, w_{tn}]$,

$$g(w_{tn}) - g(w_{t,n+1}) \geq \xi\,(w_{tn} - w_{t,n+1}) \Rightarrow g(w_{t,n+1}) - \xi w_{t,n+1} \leq g(w_{tn}) - \xi w_{tn}. \tag{B.22}$$

Combining (B.21) with (B.22), $\hat{g}_\xi(w_{tn}) \leq 0$ and $\hat{g}_\xi(w_{t,n+1}) \geq 0$, we find $\bar{\bar{g}}_\xi(w_{tn}) \leq \bar{\bar{g}}_\xi(w_{t,n+1})$, which we can write as

$$\frac{\alpha A(w_{t,n+1})}{\alpha + A(w_{t,n+1})} - \frac{\alpha A(w_{tn})}{\alpha + A(w_{tn})} + \xi \left[q(w_{t,n+1})^{-\frac{1}{\gamma}} w_{t,n+1} - q(w_{tn})^{-\frac{1}{\gamma}} w_{tn} \right] \geq 0$$

$$\Leftrightarrow \frac{\alpha^2 [A(w_{t,n+1}) - A(w_{tn})]}{[a + A(w_{tn})][\alpha + A(w_{t,n+1})](w_{t,n+1} - w_{tn})} + \xi q(w_{t,n+1})^{-\frac{1}{\gamma}} + \xi \frac{q(w_{t,n+1})^{-\frac{1}{\gamma}} - q(w_{tn})^{-\frac{1}{\gamma}}}{w_{t,n+1} - w_{tn}} w_{tn} \leq 0, \tag{B.23}$$

where the second step follows by dividing by $w_{t,n+1} - w_{tn} < 0$. The limits of the three terms in the left-hand side of (B.23) when n goes to infinity are

$$\lim_{n\to\infty} \frac{\alpha^2 [A(w_{t,n+1}) - A(w_{tn})]}{[a + A(w_{tn})][\alpha + A(w_{t,n+1})](w_{t,n+1} - w_{tn})} = \alpha^2 \lim_{w_t \to 0} \frac{A'(w_t) w_t^2}{\gamma^2}$$

$$= \alpha^2 \lim_{w_t \to 0} \frac{R'(w_t) w_t - A(w_t) w_t}{\gamma^2} = -\frac{\alpha^2}{\gamma},$$

$$\lim_{n\to\infty} \xi q(w_{t,n+1})^{-\frac{1}{\gamma}} = \xi \lim_{w_t \to 0} q(w_t)^{-\frac{1}{\gamma}},$$

$$\lim_{n\to\infty} \xi \frac{q(w_{t,n+1})^{-\frac{1}{\gamma}} - q(w_{tn})^{-\frac{1}{\gamma}}}{w_{t,n+1} - w_{tn}} w_{tn} = \xi \lim_{w_t \to 0} Q'(w_t) w_t = \xi \lim_{w_t \to 0} Q(w_t) \frac{R(w_t) - \gamma}{\gamma} = 0,$$

where the first limit follows by differentiating $R(w_t) = A(w_t) w_t$ and using $\lim_{w_t \to 0} R'(w_t) w_t = 0$ (shown in Lemma A.1) and $\lim_{w_t \to 0} A(w_t) w_t = \gamma$, and the third limit follows from (A.25) and $\lim_{w_t \to 0} R(w_t) = \gamma$. Hence, taking the limit in (B.23) when n goes to infinity, we find

$$\xi \lim_{w_t \to 0} q(w_t)^{-\frac{1}{\gamma}} \leq \frac{\alpha^2}{\gamma}.$$

Proceeding similarly for the sequence $\{\hat{w}_{tn}\}_{n \in \mathbb{N}}$, we find

$$\hat{\xi} \lim_{w_t \to 0} q(w_t)^{-\frac{1}{\gamma}} \geq \frac{\alpha^2}{\gamma},$$

which yields a contradiction since $\hat{\xi} > \xi$.[16] Suppose next, by contradiction, that $\lim_{w_t \to 0} g'(w_t) = \infty$, in which case $g''(w_t) < 0$ for small w_t. The mean-value theorem implies that there exists $\hat{w}_t \in (0, w_t)$ such that

$$r g'(w_t) - r \frac{g(w_t) + \frac{\alpha}{r}}{w_t} = r w_t g''(\hat{w}_t). \tag{B.24}$$

[16] The argument establishing that $g'(w_t)$ converges to a finite limit when w_t goes to zero can be suitably adapted in the proof of Lemma A.8 to show that the asymptotic behavior of $R(w_t)$ for w_t close to zero in the case $\gamma < K$ has the conjectured form.

Substituting into (B.18), and using $\lim_{w_t \to 0} q(w_t)^{-\frac{1}{\gamma}} \in (0, \infty)$ and $g''(w_t) < 0$ for small w_t, we find that the left-hand side converges to minus infinity when w_t goes to zero. This yields a contradiction because the right-hand side converges to a finite limit. This completes the proof of the lemma for $\gamma > K$.

Consider next the case $\gamma < K$. Since $\lim_{w_t \to 0} q(w_t)^{-\frac{1}{\gamma}} = \lim_{w_t \to 0} Q(w_t) = 0$, (B.19) rules out that $g'(w_t)$ converges to a finite limit. Moreover, the same argument as in the case $\gamma > K$ rules out that $g'(w_t)$ has no limit. Therefore, $g'(w_t) \geq 0$ converges to infinity. ∎

Equations (4.12) and (4.14) imply that the derivative of the absolute value of expected excess returns has the same sign as

$$\alpha u^\top \Sigma u \frac{[g''(w_t)A(w_t) + g'(w_t)A'(w_t)][\alpha + A(w_t)] - 2g'(w_t)A(w_t)A'(w_t)}{[\alpha + A(w_t)]^3} + \frac{\alpha A'(w_t)}{[\alpha + A(w_t)]^2}$$

$$\alpha u^\top \Sigma u \left[\frac{g''(w_t)A(w_t)}{[\alpha + A(w_t)]^2} + \frac{g'(w_t)A'(w_t)[\alpha - A(w_t)]}{[\alpha + A(w_t)]^3}\right] + \frac{\alpha A'(w_t)}{[\alpha + A(w_t)]^2}. \tag{B.25}$$

For $\gamma > K$, Lemma B.6 shows that $\lim_{w_t \to 0} g'(w_t) = \frac{\alpha^2 \gamma}{(\rho - r)z\left[\gamma + \frac{\gamma}{z}\left(\frac{r}{\rho - r} + 1\right) - 1\right]}$ and $\lim_{w_t \to 0} g''(w_t)w_t = 0$. Combining with $\lim_{w_t \to 0} A(w_t)w_t = \gamma$ and $\lim_{w_t \to 0} A'(w_t)w_t^2 = -\gamma$, we find that the limit of (B.25) when w_t goes to zero is

$$\frac{\alpha u^\top \Sigma u}{\gamma} \frac{\alpha^2 \gamma}{(\rho - r)z\left[\gamma + \frac{\gamma}{z}\left(\frac{r}{\rho - r} + 1\right) - 1\right]} - \frac{\alpha}{\gamma},$$

which is positive under the condition in the proposition. For $\gamma < K$, Lemma B.6 shows that $\lim_{w_t \to 0} g'(w_t) = \infty$, in which case $g''(w_t) < 0$ for small w_t. Equations (B.24) and $g''(w_t) < 0$ imply that the left-hand side of (B.18) is the sum of the three negative terms $\frac{\alpha^2 u^\top \Sigma u}{2[\alpha + A(w_t)]^2 w_t} g''(w_t)$, $-q(w_t)^{-\frac{1}{\gamma}} g'(w_t)$, and $r g'(w_t) - r\frac{g(w_t) + \frac{\alpha}{r}}{w_t}$. Since the right-hand side is bounded for w_t close to zero, each of the three negative terms is bounded. Since $\lim_{w_t \to 0} A(w_t)w_t = \gamma$, the boundedness of the first term means that $g''(w_t)w_t$ is bounded. Combining with $\lim_{w_t \to \infty} g'(w_t) = \infty$, $\lim_{w_t \to 0} A(w_t)w_t = \gamma$ and $\lim_{w_t \to 0} A'(w_t)w_t^2 = -\gamma$, we find that the limit of (B.25) when w_t goes to zero is infinity. ∎

C Proofs of the Results in Sections 5 and 6

Proof of Proposition 5.1: We first compute the position Y_t of arbitrageurs. Using (4.10) and (B.4), we can write (4.9) as

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

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

$$\Leftrightarrow Y_t + f(w_t) u^\top \Sigma Y_t u = \frac{\alpha}{\alpha + A(w_t) - F'(w_t)} u, \tag{C.1}$$

where I is the $N \times N$ identity matrix, and the third step follows from (3.18). Equation (C.1) implies that Y_t is collinear with u. Setting $Y_t = \nu u$ in (C.1), we find

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

and so

$$Y_t = \frac{\alpha}{[\alpha + A(w_t) - F'(w_t)][1 + u^\top \Sigma u f(w_t)]} u = \frac{\alpha}{\alpha + A(w_t) - F'(w_t) + \alpha u^\top \Sigma u g'(w_t)} u. \tag{C.2}$$

Equations (4.6) and (C.2) imply

$$\frac{\partial X_{nt}}{\partial u_n} = -\frac{\partial Y_{nt}}{\partial u_n} = -\frac{\alpha}{\alpha + A(w_t) - F'(w_t) + \alpha u^\top \Sigma u g'(w_t)}. \tag{C.3}$$

Equation (4.10) implies

$$\frac{\partial S_{nt}}{\partial u_n} = g(w_t) \Sigma_{nn}. \tag{C.4}$$

Substituting (C.3) and (C.4) into (5.1), we find (5.2). Since $\lim_{w_t \to 0} A(w_t) w_t = \max\{\gamma, K\}$, $\lim_{w_t \to 0} F'(w_t) w_t = \lim_{w_t \to 0} g'(w_t) w_t = 0$ and $\lim_{w_t \to 0} g(w_t) = -\frac{\alpha}{r}$, (5.2) implies $\lim_{w_t \to 0} \lambda_{nt} = \infty$. Since $\lim_{w_t \to \infty} A(w_t) w_t = \gamma$, $\lim_{w_t \to \infty} F'(w_t) w_t = \lim_{w_t \to \infty} g'(w_t) w_t = 0$ and $\lim_{w_t \to \infty} g(w_t) = 0$, (5.2) implies $\lim_{w_t \to \infty} \lambda_{nt} = 0$.

Suppose next that hedgers are short-lived. Since $-g(w_t) > 0$, $A(w_t) > 0$ and $g'(w_t) \geq 0$, $\lambda_{nt} > 0$. To show that λ_{nt} is decreasing in w_t if $\gamma = 1$, we prove the following lemma.

Lemma C.1 ($g'(w_t)$ decreasing if $\gamma = 1$) *For the solution to (4.11) constructed in Lemma B.5, $g'(w_t)$ is decreasing if $\gamma = 1$.*

Proof: If $g'(w_t)$ is not decreasing, then there exist $m_1 < m_2$ such that $g'(m_1) \leq g'(m_2)$. Consider first the case where $g'(m_1) \geq \lim_{w_t \to 0} g'(w_t)$, and recall from Lemma B.6 that $\lim_{w_t \to 0} g'(w_t)$ is positive for $\gamma > K$ and infinity for $\gamma < K$. Recall also from the proof of Lemma B.1 that $\lim_{w_t \to \infty} g'(w_t) w_t = 0$, and hence $\lim_{w_t \to \infty} g'(w_t) = 0$. If $g'(m_1) < g'(m_2)$, then we can choose $\xi > g'(m_1)$ and $(\hat{m}_2, \hat{\hat{m}}_2)$ such that $\hat{m}_2 < m_2 < \hat{\hat{m}}_2$, $g'(\hat{m}_2) = g'(\hat{\hat{m}}_2) = \xi$, $g''(\hat{m}_2) > 0$ and $g''(\hat{\hat{m}}_2) < 0$, and $g'(w_t) > \xi$ for all $w_t \in (\hat{m}_2, \hat{\hat{m}}_2)$. If $g'(m_1) = g'(m_2)$, then we can choose $\xi \geq g'(m_1)$ and $(\hat{m}_2, \hat{\hat{m}}_2)$ such that $\hat{m}_2 < m_2 < \hat{\hat{m}}_2$ and the remaining inequalities are weak. Proceeding as in the proof of Lemma B.6, we find $\bar{\bar{g}}_\xi(\hat{\hat{m}}_2) \leq \bar{\bar{g}}_\xi(\hat{m}_2)$. The derivative of $\bar{\bar{g}}_\xi(w_t)$ is

$$\bar{\bar{g}}'_\xi(w_t) = \xi \left[-\frac{1}{\gamma} q(w_t)^{-\frac{1}{\gamma}-1} q'(w_t) w_t + q(w_t)^{-\frac{1}{\gamma}} \right] + \frac{\alpha^2 A'(w_t)}{[\alpha + A(w_t)]^2}. \tag{C.5}$$

Since for $\gamma = 1$, $A(w_t) = \frac{1}{w_t}$ and $q(w_t) = \frac{1}{\rho}$, we can write (C.5) as

$$\bar{\bar{g}}'_\xi(w_t) = \xi \rho - \frac{\alpha^2}{(\alpha w_t + 1)^2}. \tag{C.6}$$

Lemma B.6 implies that for $\gamma = 1$, $\lim_{w_t \to 0} g'(w_t) = \frac{\alpha^2}{\rho}$. Since $\xi \geq g'(m_1) \geq \lim_{w_t \to 0} g'(w_t) = \frac{\alpha^2}{\rho}$, (C.6) implies $\bar{\bar{g}}'_\xi(w_t) > 0$ for all $w_t \in (0, \infty)$. This contradicts $\bar{\bar{g}}_\xi(\hat{\hat{m}}_2) \leq \bar{\bar{g}}_\xi(\hat{m}_2)$.

Consider next the case where $g'(m_1) < \lim_{w_t \to 0} g'(w_t)$. If $g'(m_1) < g'(m_2)$, then we can choose $\xi > g'(m_1)$ and $(\hat{m}_i, \hat{\hat{m}}_i)$ for $i = 1, 2$ such that $\hat{m}_1 < m_1 < \hat{\hat{m}}_1 \leq \hat{m}_2 < m_2 < \hat{\hat{m}}_2$, $g'(\hat{m}_i) = g'(\hat{\hat{m}}_i) = \xi$ for $i = 1, 2$, $g''(\hat{m}_1) < 0$ and $g''(\hat{\hat{m}}_1) > 0$, $g''(\hat{m}_2) > 0$ and $g''(\hat{\hat{m}}_2) < 0$, $g'(w_t) < \xi$ for all $w_t \in (\hat{m}_1, \hat{\hat{m}}_1)$, and $g'(w_t) > \xi$ for all $w_t \in (\hat{m}_2, \hat{\hat{m}}_2)$. If $g'(m_1) = g'(m_2)$, then we can choose $\xi > 0$ and $(\hat{m}_i, \hat{\hat{m}}_i)$ for $i = 1, 2$ such that $\hat{m}_1 < m_1 < \hat{\hat{m}}_1 \leq \hat{m}_2 < m_2 < \hat{\hat{m}}_2$ and the remaining inequalities are weak. Proceeding as in the case $g'(m_1) \geq \lim_{w_t \to 0} g'(w_t)$, we find $\bar{\bar{g}}_\xi(\hat{\hat{m}}_1) \geq \bar{\bar{g}}_\xi(\hat{m}_1)$ and $\bar{\bar{g}}_\xi(\hat{\hat{m}}_2) \leq \bar{\bar{g}}_\xi(\hat{m}_2)$. For $\gamma = 1$, (C.6) implies that for any $\xi > 0$, $\bar{\bar{g}}'_\xi(w_t)$ is increasing and becomes positive for large w_t. Hence, $\bar{\bar{g}}'_\xi(w_t)$ is either increasing, or decreasing and then increasing. Since, $\bar{\bar{g}}_\xi(\hat{\hat{m}}_2) \leq \bar{\bar{g}}_\xi(\hat{m}_2)$, $\bar{\bar{g}}'_\xi(w_t)$ is decreasing in at least part of $(\hat{m}_2, \hat{\hat{m}}_2)$. Hence, it is decreasing in all of $(\hat{m}_1, \hat{\hat{m}}_1)$, which contradicts $\bar{\bar{g}}_\xi(\hat{\hat{m}}_1) \geq \bar{\bar{g}}_\xi(\hat{m}_1)$. ∎

Since $-g(w_t)$ and $1 + \frac{A(w_t)}{\alpha} + u^\top \Sigma u g'(w_t)$ are positive, and $-g(w_t)$ is decreasing, λ_{nt} is decreasing if $A(w_t) + \alpha u^\top \Sigma u g'(w_t)$ is decreasing. The latter function is decreasing for $\gamma = 1$ because $A(w_t)$ is decreasing for all γ and Lemma C.1 shows that $g'(w_t)$ is decreasing for $\gamma = 1$. ∎

Proof of Corollary 5.1: We set

$$\lambda_{nt} = -g(w_t)\left(1 + \frac{A(w_t) - F'(w_t)}{\alpha} + g'(w_t)u^\top \Sigma u\right)\Sigma_{nn} \equiv L(w_t)\Sigma_{nn}. \tag{C.7}$$

Using (C.7) and Ito's lemma, we find

$$\text{Cov}_t(d\Lambda_t, dR_t) = L'(w_t)\frac{\sum_{n'=1}^N \Sigma_{n'n'}}{N}\text{Cov}_t(dw_t, dR_t), \tag{C.8}$$

$$\text{Cov}_t(d\Lambda_t, d\lambda_{nt}) = \left(L'(w_t)\right)^2 \frac{\Sigma_{nn}\sum_{n'=1}^N \Sigma_{n'n'}}{N}\text{Var}_t(dw_t), \tag{C.9}$$

$$\text{Cov}_t(d(u^\top dR_t), d\lambda_{nt}) = L'(w_t)\Sigma_{nn}u^\top \text{Cov}_t(dw_t, dR_t). \tag{C.10}$$

The diffusion matrix of the return vector dR_t is

$$\begin{aligned}
(\sigma_{St} + \sigma)^\top &= \left(\frac{\alpha}{\alpha + A(w_t) - F'(w_t)}\sigma u S'(w_t)^\top + \sigma\right)^\top \\
&= \left(\frac{\alpha g'(w_t)}{\alpha + A(w_t) - F'(w_t)}\sigma u u^\top \Sigma + \sigma\right)^\top,
\end{aligned} \tag{C.11}$$

where the first step follows from (B.4) and the second from (4.10). The covariance between wealth and the return vector dR_t is

$$\begin{aligned}
\text{Cov}_t(dw_t, dR_t) &= (\sigma_{St} + \sigma)^\top \sigma_{wt} \\
&= \left(\frac{\alpha g'(w_t)}{\alpha + A(w_t) - F'(w_t)}\sigma u u^\top \Sigma + \sigma\right)^\top \frac{\alpha}{\alpha + A(w_t) - F'(w_t)}\sigma u \\
&= \frac{\alpha}{\alpha + A(w_t) - F'(w_t)}\left[u^\top \Sigma u f(w_t) + 1\right]\Sigma u,
\end{aligned} \tag{C.12}$$

where the second step follows from (A.80) and (C.11). Part (i) of the corollary follows by substituting (C.12) into (C.8). The proportionality coefficient is

$$C^\Lambda(w_t) = L'(w_t)\frac{\alpha \sum_{n'=1}^N \Sigma_{n'n'}}{N[\alpha + A(w_t) - F'(w_t)]}\left[u^\top \Sigma u f(w_t) + 1\right], \tag{C.13}$$

and is negative when hedgers are short-lived and arbitrageurs have logarithmic preferences because $L(w_t)$ is decreasing in w_t (Proposition 5.1). Part (ii) of the corollary follows from (C.9). The proportionality coefficient is positive regardless of the shape of $L(w_t)$. Part (iii) of the corollary follows by substituting (C.12) into (C.10). The proportionality coefficient is negative when hedgers are short-lived and arbitrageurs have logarithmic preferences because $L(w_t)$ is decreasing in w_t. ∎

Proof of Corollary 5.2: The proportionality result follows from (4.12), (C.8), and (C.12). These equations imply that the proportionality coefficient is

$$\Pi^\Lambda(w_t) = \frac{A(w_t)}{L'(w_t)\frac{\sum_{n'=1}^N \Sigma_{n'n'}}{N}}. \tag{C.14}$$

This coefficient is negative when hedgers are short-lived and arbitrageurs have logarithmic preferences because $L(w_t)$ is decreasing in w_t. ∎

Proof of Proposition 5.2: Using (C.7) to compute $L'(w_t)$, we can write (C.13) as

$$\begin{aligned} C^\Lambda(w_t) =& \frac{\left[-g'(w_t)\left(1+\frac{A(w_t)-F'(w_t)}{\alpha}+g'(w_t)u^\top\Sigma u\right)-g(w_t)\left(\frac{A'(w_t)-F''(w_t)}{\alpha}+g''(w_t)u^\top\Sigma u\right)\right]}{\alpha+A(w_t)-F'(w_t)} \\ & \times \frac{\alpha\sum_{n'=1}^N \Sigma_{n'n'}}{N}\left[u^\top\Sigma u f(w_t)+1\right], \end{aligned} \tag{C.15}$$

and $\Pi^\Lambda(w_t)$ as

$$\begin{aligned} \Pi^\Lambda(w_t) =& \frac{A(w_t)}{\left[-g'(w_t)\left(1+\frac{A(w_t)-F'(w_t)}{\alpha}+g'(w_t)u^\top\Sigma u\right)-g(w_t)\left(\frac{A'(w_t)-F''(w_t)}{\alpha}+g''(w_t)u^\top\Sigma u\right)\right]} \\ & \times \frac{\sum_{n'=1}^N \Sigma_{n'n'}}{N}. \end{aligned} \tag{C.16}$$

We denote the numerator and denominator of the fraction in the first line of (C.15) by $N_C(w_t)$ and $D_C(w_t)$, respectively. We denote the numerator and denominator of the fraction in the first line of (C.16) by $N_\Pi(w_t)$ and $D_\Pi(w_t)$, respectively, noting that $N_C(w_t) = D_\Pi(w_t)$.

Since $\lim_{w_t\to 0} A(w_t)w_t = \max\{\gamma, K\}$ and $\lim_{w_t\to 0} F'(w_t)w_t = 0$, $\lim_{w_t\to 0} D_C(w_t)w_t = \max\{\gamma, K\}$. Since, in addition, $\lim_{w_t\to 0} g(w_t) = -\frac{\alpha}{r}$, $\lim_{w_t\to 0} g'(w_t)w_t = \lim_{w_t\to 0} g''(w_t)w_t^2 = \lim_{w_t\to 0} F''(w_t)w_t^2 = 0$ and $\lim_{w_t\to 0} A'(w_t)w_t^2 = -\max\{\gamma, K\}$, $\lim_{w_t\to 0} N_C(w_t)w_t^2 = -\frac{\max\{\gamma,K\}}{r}$. Hence $\lim_{w_t\to 0} N_C(w_t)w_t = -\infty$. Since, in addition, $\lim_{w_t\to 0} f(w_t) = 0$, (C.15) implies that $\lim_{w_t\to 0} C^\Lambda(w_t) = -\infty$.

Since $\lim_{w_t\to 0} A(w_t)w_t = \max\{\gamma, K\}$, $\lim_{w_t\to 0} N_\Pi(w_t)w_t = \max\{\gamma, K\}$. Since, in addition, $D_\Pi(w_t) = N_C(w_t)$, $\lim_{w_t\to 0} D_\Pi(w_t)w_t = -\infty$. Therefore, (C.16) implies that $\lim_{w_t\to 0} \Pi^\Lambda(w_t) = 0$.

Since $\lim_{w_t\to\infty} A(w_t) = \lim_{w_t\to\infty} F'(w_t) = 0$ (because the limits of these functions multiplied by w_t are γ and zero, respectively), $\lim_{w_t\to\infty} D_C(w_t) = \alpha$. Since, in addition, $\lim_{w_t\to\infty} g(w_t) =$

$\lim_{w_t\to\infty} g''(w_t) = 0$ and $\lim_{w_t\to\infty} A'(w_t) = \lim_{w_t\to\infty} F''(w_t) = \lim_{w_t\to\infty} g'(w_t) = 0$ (because the limits of the first two functions multiplied by w_t^2 are $-\gamma$ and zero, respectively, and the limit of the third function multiplied by w_t is zero), $\lim_{w_t\to\infty} N_C(w_t) = 0$. Since, in addition, $\lim_{w_t\to\infty} f(w_t) = 0$, (C.15) implies that $\lim_{w_t\to\infty} C^\Lambda(w_t) = 0$.

We finally determine $\lim_{w_t\to\infty} \Pi^\Lambda(w_t)$ when hedgers are short-lived. Since $\lim_{w_t\to\infty} A(w_t)w_t = \gamma$, $\lim_{w_t\to\infty} N_\Pi(w_t)w_t = \gamma$. We next derive $\lim_{w_t\to\infty} D_\Pi(w_t)w_t$ assuming that $g''(w_t)w_t^2$ has a (finite or infinite) limit when w_t goes to infinity. Since $\lim_{w_t\to\infty} g(w_t)$ is finite, $\lim_{w_t\to\infty} g''(w_t)w_t^2 = 0$. Since $\lim_{w_t\to\infty} A(w_t) = \lim_{w_t\to\infty} A'(w_t)w_t = \lim_{w_t\to\infty} g(w_t) = \lim_{w_t\to\infty} g'(w_t)w_t = \lim_{w_t\to\infty} g'(w_t) = \lim_{w_t\to\infty} g''(w_t)w_t = 0$, $\lim_{w_t\to\infty} D_\Pi(w_t)w_t = 0$. Therefore, (C.16) implies that $\lim_{w_t\to\infty} \Pi^\Lambda(w_t)$ is plus or minus infinity. To show that the limit is $-\infty$, we note that the largest-order term in $-g'(w_t)\left(1 + \frac{A(w_t)}{\alpha} + g'(w_t)u^\top\Sigma u\right)$ is $-g'(w_t)$, which is non-positive because $g(w_t)$ is increasing, and that the largest-order term in $-g(w_t)\left(\frac{A'(w_t)}{\alpha} + g''(w_t)u^\top\Sigma u\right)$ is $-g(w_t)\frac{A'(w_t)}{\alpha}$, which is negative because $g(w_t) < 0$ and $A'(w_t) < 0$. ∎

Proof of Proposition 6.1: We first derive the positions of hedgers and arbitrageurs, as well as asset prices, in an equilibrium with positive supply. We next derive the Bellman equations of hedgers and arbitrageurs. We finally show the equivalence with a zero-supply equilibrium.

Positive supply does not change the asset demands (3.6) and (3.13) of hedgers and arbitrageurs. We write these demands in terms of the long-maturity assets, using the mapping derived in Lemma 4.1. (That is, we use (4.7) to replace $\bar{D} - \pi_t$ by $\mu_{St} + \bar{D} - rS_t$, and (4.8) and (4.9) to replace (x_t, y_t) by (X_t, Y_t).) Equations (3.7) and (3.13) become

$$X_t = \frac{\left[(\sigma_{St}+\sigma)^\top(\sigma_{St}+\sigma)\right]^{-1}(\mu_{St}+\bar{D}-rS_t)}{\alpha} - (\sigma_{St}+\sigma)^{-1}\sigma u - \frac{F'(w_t)Y_t}{\alpha}, \tag{C.17}$$

$$Y_t = \frac{\left[(\sigma_{St}+\sigma)^\top(\sigma_{St}+\sigma)\right]^{-1}(\mu_{St}+\bar{D}-rS_t)}{A(w_t)}, \tag{C.18}$$

respectively. Using the market-clearing equation

$$X_t + Y_t = s, \tag{C.19}$$

we can write (C.17) as

$$X_t = \frac{\left[(\sigma_{St}+\sigma)^\top(\sigma_{St}+\sigma)\right]^{-1}(\mu_{St}+\bar{D}-rS_t) - \alpha(\sigma_{St}+\sigma)^{-1}\sigma u - F'(w_t)s}{\alpha - F'(w_t)}. \tag{C.20}$$

Substituting X_t and Y_t from (C.20) and (C.18), respectively, into (C.19), we find that expected excess returns are

$$\mu_{St} + \bar{D} - rS_t = \frac{\alpha A(w_t)}{\alpha + A(w_t) - F'(w_t)} (\sigma_{St} + \sigma)^\top b_t, \tag{C.21}$$

where

$$b_t \equiv (\sigma_{St} + \sigma)s + \sigma u. \tag{C.22}$$

Substituting $\mu_{St} + \bar{D} - rS_t$ from (C.21) into (C.18), we find that the position of arbitrageurs is

$$Y_t = \frac{\alpha}{\alpha + A(w_t) - F'(w_t)} (\sigma_{St} + \sigma)^{-1} b_t. \tag{C.23}$$

Substituting c_t from (A.14), $\mu_{St} + \bar{D} - rS_t$ from (C.21), and Y_t from (C.23) into (B.2), we find that the dynamics of arbitrageur wealth are given by (A.78) with

$$\mu_{wt} = \left(r - q(w_t)^{-\frac{1}{\gamma}}\right) w_t + \frac{\alpha^2 A(w_t)}{[\alpha + A(w_t) - F'(w_t)]^2} b_t^\top b_t, \tag{C.24}$$

$$\sigma_{wt} = \frac{\alpha}{\alpha + A(w_t) - F'(w_t)} b_t. \tag{C.25}$$

Using (C.24) and (C.25), we find the following counterparts of (B.3) and (B.4):

$$\mu_{St} = \left(r - q(w_t)^{-\frac{1}{\gamma}}\right) S'(w_t) w_t + \frac{\alpha^2}{[\alpha + A(w_t) - F'(w_t)]^2} b_t^\top b_t \left(A(w_t) S'(w_t) + \frac{1}{2} S''(w_t)\right), \tag{C.26}$$

$$\sigma_{St} = \frac{\alpha}{\alpha + A(w_t) - F'(w_t)} b_t S'(w_t)^\top. \tag{C.27}$$

Substituting σ_{St} from (C.27) into (C.22) and solving for b_t, we find

$$b_t = \frac{\sigma(s + u)}{1 - \frac{\alpha}{\alpha + A(w_t) - F'(w_t)} S'(w_t)^\top s}. \tag{C.28}$$

Substituting (μ_{St}, σ_{St}) from (C.26) and (C.27) into (C.21), we find

$$\begin{aligned}
&\left(r - q(w_t)^{-\frac{1}{\gamma}}\right) S'(w_t) w_t + \frac{\alpha^2}{2[\alpha + A(w_t) - F'(w_t)]^2} b_t^\top b_t S''(w_t) + \bar{D} - rS(w_t) \\
&= \frac{\alpha A(w_t)}{\alpha + A(w_t) - F'(w_t)} \sigma^\top b_t \\
&\Leftrightarrow \left(r - q(w_t)^{-\frac{1}{\gamma}}\right) S'(w_t) w_t + \frac{\alpha^2 (s + u)^\top \Sigma (s + u)}{2\left[\alpha + A(w_t) - F'(w_t) - \alpha S'(w_t)^\top s\right]^2} S''(w_t) + \bar{D} - rS(w_t) \\
&= \frac{\alpha A(w_t)}{\alpha + A(w_t) - F'(w_t) - \alpha S'(w_t)^\top s} \Sigma (s + u),
\end{aligned} \tag{C.29}$$

where the second step follows from (C.28). The same argument as in the proof of Proposition 4.1 implies that $S(w_t)$ must have the form in (4.10). Substituting (4.10) into (C.29), we find that $g(w_t)$ solves the ODE

$$\frac{\alpha^2(s+u)^\top\Sigma(s+u)}{2\left[\alpha+A(w_t)-F'(w_t)-\alpha g'(w_t)(s+u)^\top\Sigma s\right]^2}g''(w_t)+\left(r-q(w_t)^{-\frac{1}{\gamma}}\right)g'(w_t)w_t-rg(w_t)$$
$$=\frac{\alpha A(w_t)}{\alpha+A(w_t)-F'(w_t)-\alpha g'(w_t)(s+u)^\top\Sigma s}. \tag{C.30}$$

Substituting S_t from (4.10) into (C.28), we can write b_t as

$$b_t=\frac{\sigma(s+u)}{1-\frac{\alpha}{\alpha+A(w_t)-F'(w_t)}g'(w_t)(s+u)^\top\Sigma s}. \tag{C.31}$$

Substituting b_t from (C.31) into (C.21) and (C.23), we can write $\mu_{St}+\bar{D}-rS_t$ and Y_t as

$$\mu_{St}+\bar{D}-rS_t=\frac{\alpha A(w_t)}{\alpha+A(w_t)-F'(w_t)-\alpha g'(w_t)(s+u)^\top\Sigma s}(\sigma_{St}+\sigma)^\top\sigma(s+u), \tag{C.32}$$

$$Y_t=\frac{\alpha}{\alpha+A(w_t)-F'(w_t)-\alpha g'(w_t)(s+u)^\top\Sigma s}(\sigma_{St}+\sigma)^{-1}\sigma(s+u), \tag{C.33}$$

respectively. This completes the derivation of positions and prices: the position Y_t of arbitrageurs is given by (C.33), the position X_t of hedgers is given by $s-Y_t$, and asset prices are given by (4.10), where $g(w_t)$ solves the ODE (C.30).

We next derive the Bellman equation of arbitrageurs. To do that, we write (A.12) in terms of the long-maturity assets, using (4.7) to replace $\bar{D}-\pi_t$ by $\mu_{St}+\bar{D}-rS_t$, and (4.8) and (4.9) to replace (x_t,y_t) by (X_t,Y_t). We also note that the maximum in (A.12) is achieved for $\hat{c}_t$ given by (3.11) and for $\hat{y}_t=y_t\frac{\hat{w}_t}{w_t}$, and we substitute c_t from (A.14), $\mu_{St}+\bar{D}-rS_t$ from (C.32), and Y_t from (C.33). This yields the following counterpart of (A.16):

$$\begin{aligned}
\rho q(w_t) &= q(w_t)^{1-\frac{1}{\gamma}}\\
&+\left(q'(w_t)+\frac{(1-\gamma)q(w_t)}{w_t}\right)\left(rw_t-q(w_t)^{-\frac{1}{\gamma}}w_t+\frac{\alpha^2(s+u)^\top\Sigma(s+u)A(w_t)}{[\alpha+A(w_t)-F'(w_t)-\alpha g'(w_t)(s+u)^\top\Sigma s]^2}\right)\\
&+\frac{1}{2}\left(q''(w_t)-\frac{\gamma(1-\gamma)q(w_t)}{w_t^2}+\frac{2(1-\gamma)q'(w_t)}{w_t}\right)\frac{\alpha^2(s+u)^\top\Sigma(s+u)}{[\alpha+A(w_t)-F'(w_t)-\alpha g'(w_t)(s+u)^\top\Sigma s]^2}.
\end{aligned} \tag{C.34}$$

Proceeding as in the proof of Proposition 3.4, we find the following counterpart of (3.21):

$$1=\frac{q(w_t)^{-\frac{1}{\gamma}}-r}{\rho-r}A(w_t)w_t-\frac{z_s\left(A'(w_t)+A(w_t)^2\right)}{[\alpha+A(w_t)-F'(w_t)-\alpha g'(w_t)(s+u)^\top\Sigma s]^2}, \tag{C.35}$$

where

$$z_s \equiv \frac{\alpha^2(s+u)^\top\Sigma(s+u)}{2(\rho-r)}. \tag{C.36}$$

We next derive the Bellman equation of hedgers. Using (4.7) to replace $\bar{D}-\pi_t$ by $\mu_{St}+\bar{D}-rS_t$, and (4.8) and (4.9) to replace (x_t, y_t) by (X_t, Y_t), we find the following counterpart of (A.21):

$$\begin{aligned}
&\bar{\rho} = r + rF(w_t) - r\log(r) - \alpha u^\top \bar{D} - \frac{1}{2}\alpha^2\left[X_t^\top(\sigma_{St}+\sigma)^\top - u^\top\sigma^\top\right][(\sigma_{St}+\sigma)X_t + u\sigma] \\
&- F'(w_t)\left[rw_t - c_t + Y_t^\top(\mu_{St}+\bar{D}-rS_t)\right] \\
&- \frac{1}{2}\left[F''(w_t) - F'(w_t)^2\right]Y_t^\top(\sigma_{St}+\sigma)^\top(\sigma_{St}+\sigma)Y_t + \alpha F'(w_t)u^\top\sigma^\top(\sigma_{St}+\sigma)Y_t \\
&\Leftrightarrow \bar{\rho} = r + rF(w_t) - r\log(r) - \alpha u^\top\bar{D} - \frac{1}{2}\alpha^2\left[b_t - (\sigma_{St}+\sigma)Y_t - 2u^\top\sigma^\top\right][b_t - (\sigma_{St}+\sigma)Y_t] \\
&- F'(w_t)\left[rw_t - c_t + Y_t^\top(\mu_{St}+\bar{D}-rS_t)\right] \\
&- \frac{1}{2}\left[F''(w_t) - F'(w_t)^2\right]Y_t^\top(\sigma_{St}+\sigma)^\top(\sigma_{St}+\sigma)Y_t + \alpha F'(w_t)u^\top\sigma^\top(\sigma_{St}+\sigma)Y_t,
\end{aligned} \tag{C.37}$$

where the second step follows from (C.19) and (C.22). Using (C.23), we find

$$b_t - (\sigma_{St}+\sigma)Y_t = \frac{A(w_t) - F'(w_t)}{\alpha + A(w_t) - F'(w_t)}b_t = \frac{[A(w_t) - F'(w_t)]\sigma(s+u)}{\alpha + A(w_t) - F'(w_t) - \alpha g'(w_t)(s+u)^\top\Sigma s}, \tag{C.38}$$

where the second step follows from (C.31). Substituting c_t from (A.14), $\mu_{St}+\bar{D}-rS_t$ from (C.32), Y_t from (C.33), and $b_t - (\sigma_{St}+\sigma)Y_t$ from (C.38), we can write (C.37) as

$$\begin{aligned}
\bar{\rho} =& r + rF(w_t) - r\log(r) - \alpha u^\top\bar{D} - \left(r - q(w_t)^{-\frac{1}{\gamma}}\right)F'(w_t)w_t \\
&- \frac{\alpha^2(s+u)^\top\Sigma(s+u)\left\{F''(w_t) + A(w_t)^2\right\}}{[\alpha + A(w_t) - F'(w_t) - \alpha g'(w_t)(s+u)^\top\Sigma s]^2} + \frac{\alpha^2 u^\top\Sigma(s+u)A(w_t)}{\alpha + A(w_t) - F'(w_t) - \alpha g'(w_t)(s+u)^\top\Sigma s}.
\end{aligned} \tag{C.39}$$

The equilibrium is characterized by the ODEs (3.14), (C.30), (C.35) and (C.39). To show the equivalence with a zero-supply equilibrium, we define the function

$$\hat{F}(w_t) \equiv F(w_t) - \alpha g(w_t)(s+u)^\top\Sigma s. \tag{C.40}$$

Using (C.40), we can write (C.30) and (C.35) as

$$\frac{\alpha^2(s+u)^\top\Sigma(s+u)}{2\left[\alpha+A(w_t)-\hat{F}'(w_t)\right]^2}g''(w_t)+\left(r-q(w_t)^{-\frac{1}{\gamma}}\right)g'(w_t)w_t-rg(w_t)=\frac{\alpha A(w_t)}{\alpha+A(w_t)-\hat{F}'(w_t)}, \tag{C.41}$$

$$1=\frac{q(w_t)^{-\frac{1}{\gamma}}-r}{\rho-r}A(w_t)w_t-\frac{z_s\left(A'(w_t)+A(w_t)^2\right)}{[\alpha+A(w_t)-\hat{F}'(w_t)]^2}, \tag{C.42}$$

respectively. We next multiply (C.30) by $\alpha(s+u)^\top\Sigma s$ and add it to (C.39). Using (C.40), we can write the resulting equation as

$$\begin{aligned}
\bar{\rho}&=r+r\hat{F}(w_t)-r\log(r)-\alpha u^\top\bar{D}-\left(r-q(w_t)^{-\frac{1}{\gamma}}\right)\hat{F}'(w_t)w_t\\
&-\frac{\alpha^2(s+u)^\top\Sigma(s+u)\left\{\hat{F}''(w_t)+A(w_t)^2\right\}}{[\alpha+A(w_t)-F'(w_t)-\alpha g'(w_t)(s+u)^\top\Sigma s]^2}+\frac{\alpha^2(s+u)^\top\Sigma(s+u)A(w_t)}{\alpha+A(w_t)-F'(w_t)-\alpha g'(w_t)(s+u)^\top\Sigma s}\\
\Leftrightarrow 1&=\frac{r\hat{F}(w_t)-r\log(r)-\alpha u^\top\bar{D}+\rho-\bar{\rho}}{\rho-r}+\frac{q(w_t)^{-\frac{1}{\gamma}}-r}{\rho-r}\hat{F}'(w_t)w_t\\
&-\frac{z_s\left\{\hat{F}''(w_t)-A(w_t)\left[2\alpha+A(w_t)-2\hat{F}'(w_t)\right]\right\}}{\left[\alpha+A(w_t)-\hat{F}'(w_t)\right]^2}.
\end{aligned} \tag{C.43}$$

Equations (C.41), (C.42) and (C.43) are identical to (4.11), (3.21) and (3.22), respectively, except that $F(w_t)$ is replaced by $\hat{F}(w_t)$ and u is replaced by $s+u$. Hence, a solution $(q(w_t), A(w_t), \hat{F}(w_t), g(w_t))$ to the system of (3.14), (C.41), (C.42) and (C.43), coincides with a solution $(q(w_t), A(w_t), F(w_t), g(w_t))$ to the system of (3.14), (3.21), (3.22) and (4.11) provided that we replace u by $s+u$ in the latter system. Since the function $g(w_t)$ is identical in the two cases, the price is the same in the positive-supply equilibrium and in a zero-supply equilibrium in which u is replaced by $s+u$. The position Y_t of arbitrageurs is also the same across the two equilibria because of (C.33) and $\hat{F}(w_t)=F(w_t)$. Since Y_t is the same, the market-clearing equation (3.15) implies that the position X_t of hedgers in the zero-supply equilibrium is equal to that in the positive-supply equilibrium minus s. Since positions and prices are the same across the two equilibria, the exposures of hedgers and arbitrageurs to the Brownian shocks are also the same. ∎

Proof of Proposition 6.2: The arguments in the proof of Proposition 6.1 that concern prices, positions and the arbitrageurs' Bellman equation remain valid for short-lived hedgers, provided that we set $F(w_t) = 0$. Hence, $S(w_t)$ has the form in (4.10) and $g(w_t)$ solves the ODE

$$\frac{\alpha^2 (s+u)^\top \Sigma (s+u)}{2\left[\alpha + A(w_t) - \alpha g'(w_t)(s+u)^\top \Sigma s\right]^2} g''(w_t) + \left(r - q(w_t)^{-\frac{1}{\gamma}}\right) g'(w_t) w_t - r g(w_t)$$
$$= \frac{\alpha A(w_t)}{\alpha + A(w_t) - \alpha g'(w_t)(s+u)^\top \Sigma s}, \tag{C.44}$$

which is obtained from (C.30) by setting $F(w_t) = 0$. Substituting μ_{St} from (C.26) into (4.2), and using (4.10), (C.30), (C.31) and $F(w_t) = 0$, we can write expected excess returns as (4.12). ∎