Submitted to Econometrica

SUPPLEMENT TO "PRICING AND LIQUIDITY IN DECENTRALIZED ASSET MARKETS"

SEMİH ÜSLÜ

APPENDIX B: OPTIMIZATION

This appendix covers the stochastic control problem that an individual investor with the reduced-form quasi-linear utility faces in the OTC market equilibrium of Section 2. I define the investor's problem and provide HJB equations and an optimality verification argument along the lines of Duffie, Gârleanu, and Pedersen (2005) and Vayanos and Weill (2008). I conclude by establishing the existence and uniqueness of the solution to the individual investor's problem taking as given the joint distribution of taste types, asset positions, and speed types.

Investor's Problem

I fix a probability space $(\Omega, \mathcal{F}, \Pr)$ and a filtration $\{\mathcal{F}_t, t \geq 0\}$ of sub-σ-algebras satisfying the usual conditions (see Protter, 2004). An investor can be of either one of the three-dimensional continuum of types denoted by $(\delta, a, \lambda) \in \mathcal{T} \equiv [\delta_L, \delta_H] \times \mathbb{R} \times [0, M]$. The arrival times of changes of taste types and of potential counterparties are counted by two independent adapted counting processes N^α and N^λ with constant intensities α and $m(\lambda, \Lambda)$, respectively. The details of these counting processes that govern idiosyncratic shocks and trade are as described in Section 2.

An investor with initial type (δ_0, a_0, λ) and initial wealth W_0 chooses a feasible trading strategy $\{a_t\}_{t \in [0,\infty)}$ and an adapted consumption and wealth process $\{(c_t, W_t)\}_{t \in [0,\infty)}$ subject to the following feasibility conditions. First, the type (δ_t, a_t, λ) must remain constant during the inter- and intra-arrival times of the counting processes N^α and N^λ. Second, when the investor is in state $(\delta, a, \lambda) \in \mathcal{T}$ and when the process N_t^α jumps, the investor transitions into the state $(\delta', a, \lambda) \in \mathcal{T}$, where the investor's new taste type, δ', is drawn according to the pdf f on $[\delta_L, \delta_H]$. Third, when the investor is in state $(\delta, a, \lambda) \in \mathcal{T}$ and when the process N_t^λ jumps, the investor transitions into the state $(\delta, a + q_t[(\delta, a, \lambda), (\delta', a', \lambda')], \lambda) \in \mathcal{T}$, where the trade quantity, $q_t[(\delta, a, \lambda), (\delta', a', \lambda')]$, is bargained with the counterparty of type (δ', a', λ') who is drawn according to the stationary joint cdf, $\Phi(\delta', a', \lambda')$, of taste types, asset positions, and speed types, with the likelihood, $\frac{m(\lambda, \lambda')}{m(\lambda, \Lambda)}$, that depends on her speed type λ'.[1]

Johns Hopkins Carey Business School, 100 International Drive, Baltimore, MD 21202.
semihuslu@jhu.edu

[1]Since investors have quasi-linear preferences, terms of trade are independent of wealth levels, as will be clear shortly.

First, I start by describing an investor's continuation utility at time t from remaining lifetime consumption. For a particular investor, the arguments of this continuation utility function are, naturally, the investor's current wealth W_t, her current type (δ_t, a_t, λ), and time t. More precisely, the continuation utility is

$$U(W_t, \delta_t, a_t, \lambda, t) = \sup_{C,a} \mathbb{E}_t \int_0^\infty e^{-rs} dC_{t+s} \tag{B.1}$$

s.t.

$$dW_t = rW_t dt - dC_t + u(\delta_t, a_t)\, dt - P_t\left[(\delta_{t-}, a_{t-}, \lambda), (\delta'_t, a'_t, \lambda'_t)\right] da_t, \tag{B.2}$$

$$da_t = \begin{cases} q_t\left[(\delta_{t-}, a_{t-}, \lambda), (\delta'_t, a'_t, \lambda'_t)\right] & \text{if } (\delta'_t, a'_t, \lambda'_t) \text{ is contacted} \\ 0 & \text{if no contact,} \end{cases}$$

where

$$\begin{aligned} &\{q_t\left[(\delta, a, \lambda), (\delta', a', \lambda')\right], P_t\left[(\delta, a, \lambda), (\delta', a', \lambda')\right]\} = \\ &\arg\max_{q,P} \Big\{\left[U(W - qP, \delta, a + q, \lambda, t) - U(W, \delta, a, \lambda, t)\right]^{\frac{1}{2}} \\ &\qquad\qquad \left[U(W' + qP, \delta', a' - q, \lambda', t) - U(W', \delta', a', \lambda', t)\right]^{\frac{1}{2}}\Big\}, \end{aligned}$$

s.t.

$$\begin{aligned} &U(W - qP, \delta, a + q, \lambda, t) \geq U(W, \delta, a, \lambda, t), \\ &U(W' + qP, \delta', a' - q, \lambda', t) \geq U(W', \delta', a', \lambda', t). \end{aligned}$$

where $\mathbb{E}_t$ denotes expectation conditional on the information at time t, $\{C_t\}_{t\in[0,\infty)}$ is a cumulative consumption process, $\{(\rho_t, a_t, \lambda)\}_{t\in[0,\infty)}$ is a $\mathcal{T}$-valued type process induced by the feasible trading strategy $\{a_t\}_{t\in[0,\infty)}$, and the benefit $u(\rho_t, a_t)$ has a similar holding benefit/cost interpretation as in Duffie et al. (2005). The difference is that I assume the holding benefit is a concave quadratic function of asset position while it is linear in Duffie et al. (2005). (B.1) and (B.2) imply that the continuation utility is linear in wealth, i.e., $U(W_t, \delta_t, a_t, \lambda, t) = W_t + J(\delta_t, a_t, \lambda, t)$, where

$$\begin{aligned} &J(\delta_t, a_t, \lambda, t) = \\ &\sup_a \mathbb{E}_t\left[\int_t^\infty e^{-r(s-t)} u(\delta_s, a_s)\, ds - e^{-r(s-t)} P_s\left[(\delta_{s-}, a_{s-}, \lambda), (\delta'_s, a'_s, \lambda'_s)\right] da_s\right]. \end{aligned} \tag{B.3}$$

Finally, to guarantee the global optimality of the trading strategy induced by (B.3), I impose the transversality condition

$$\lim_{t\to\infty} e^{-rt} J(\delta, a, \lambda, t) = 0 \tag{B.4a}$$

for all $(\delta, a, \lambda) \in \mathcal{T}$ and the condition

$$\text{(B.4b)} \qquad \mathbb{E}\left[\int_0^T \left(e^{-rs} J\left(\delta_s, a_s, \lambda, s\right)\right)^2 ds\right] < \infty$$

for any $T > 0$, for any initial investor type (δ_0, a_0, λ), any feasible trading strategy $\{a_t\}_{t\in[0,\infty)}$, and the associated type process $\{(\delta_t, a_t, \lambda)\}_{t\in[0,\infty)}$. These conditions will allow me to complete the usual verification argument for stochastic control.

HJB Equations

In order to derive J, q, and P, I focus on a particular investor and a particular time t. I let τ_α be an exponential random variable that represents the next (stopping) time at which that investor's taste type changes, let τ_λ be an exponential random variable that represents the next (stopping) time at which another investor is met, and let $\tau = \min\{\tau_\alpha, \tau_\lambda\}$. Then,

$$\text{(B.5)} \qquad \begin{aligned} &J\left(\delta_t, a_t, \lambda, t\right) \\ &= \mathbb{E}_t\left[\int_t^\tau e^{-r(s-t)} u\left(\delta_s, a_s\right) ds + e^{-r(\tau_\alpha - t)} \mathbb{I}_{\{\tau_\alpha = \tau\}} \int_{\delta_L}^{\delta_H} J(\delta', a_t, \lambda) f(\delta') d\delta'\right. \\ &+ e^{-r(\tau_\lambda - t)} \mathbb{I}_{\{\tau_\lambda = \tau\}} \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{m\left(\lambda, \lambda'\right)}{m\left(\lambda, \Lambda\right)} \{J(\delta_t, a_t + q_{\tau_\lambda}[(\delta, a, \lambda), (\delta', a', \lambda')], \lambda) \\ &\left. - q_{\tau_\lambda}\left[\left(\delta_t, a_t, \lambda\right), \left(\delta', a', \lambda'\right)\right] P_{\tau_\lambda}\left[\left(\delta_t, a_t, \lambda\right), \left(\delta', a', \lambda'\right)\right]\right\} \Phi(d\delta', da', d\lambda')\right]. \end{aligned}$$

Differentiating the both sides of (B.5) with respect to time argument t and suppressing it, I arrive at

$$\text{(B.6)} \qquad \begin{aligned} \dot{J}\left(\delta, a, \lambda\right) &= rJ(\delta, a, \lambda) - u(\delta, a) - \alpha \int_{\delta_L}^{\delta_H} [J(\delta', a, \lambda) - J(\delta, a, \lambda)] f(\delta') d\delta' \\ &- \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} m\left(\lambda, \lambda'\right) \left\{J(\delta, a + q\left[\left(\delta, a, \lambda\right), \left(\delta', a', \lambda'\right)\right], \lambda) - J(\delta, a, \lambda)\right. \\ &\qquad \left. - q\left[\left(\delta, a, \lambda\right), \left(\delta', a', \lambda'\right)\right] P\left[\left(\delta, a, \lambda\right), \left(\delta', a', \lambda'\right)\right]\right\} \Phi(d\delta', da', d\lambda'). \end{aligned}$$

In steady state, $\dot{J}\left(\delta, a, \lambda\right) = 0$ and hence (B.6) implies the HJB equation (3.1) of Section 3. After using the Nash bargaining procedure for the determination of

$q\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]$ and $P\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]$, I get the auxiliary HJB equation (3.13) of Subsection 3.3:

$$\begin{aligned}\text{(B.7)}\qquad rJ(\delta,a,\lambda) &= \delta a-\frac{1}{2}\kappa a^2+\alpha\int_{\delta_L}^{\delta_H}\left[J(\delta',a,\lambda)-J(\delta,a,\lambda)\right]f(\delta')d\delta' \\ &+\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m\left(\lambda,\lambda'\right)\frac{1}{2}\left[\max_q\ \{J(\delta,a+q,\lambda)-J(\delta,a,\lambda)\right. \\ &\left.+J(\delta',a'-q,\lambda')-J(\delta',a',\lambda')\}\right]\Phi(d\delta',da',d\lambda').\end{aligned}$$

Optimality Verification

Now, to verify the sufficiency of the HJB equation (3.1) for individual optimality, I consider any initial investor type (δ_0,a_0,λ), any feasible trading strategy $\{a_t\}_{t\in[0,\infty)}$, and the associated type process $\{(\delta_t,a_t,\lambda)\}_{t\in[0,\infty)}$. I assume, without loss of generality, the wealth process is $W_t=0$ for all $t\geq 0$. Therefore, the resulting cumulative consumption process $\{C_t^a\}_{t\in[0,\infty)}$ satisfies

$$\text{(B.8)}\qquad dC_t^a=u\left(\delta_t,a_t\right)dt-P_t\left[\left(\delta_{t-},a_{t-},\lambda\right),\left(\delta_t',a_t',\lambda_t'\right)\right]da_t.$$

At any time $T>0$,

$$\begin{aligned}\text{(B.9)}\qquad &\mathbb{E}\left[\int_0^T e^{-rs}dC_s^a+e^{-rT}J(\delta_T,a_T,\lambda)\right] \\ &=\mathbb{E}\left[\int_0^T e^{-rs}dC_s^a+J(\delta_0,a_0,\lambda)+\int_0^T d\left(e^{-rs}J(\delta_s,a_s,\lambda)\right)\right] \\ &=\mathbb{E}\left[J(\delta_0,a_0,\lambda)+\int_0^T e^{-rs}dC_s^a+\int_0^T\left(-re^{-rs}J(\delta_s,a_s,\lambda)\right)ds\right. \\ &\left.+\int_0^T e^{-rs}d\left(J(\delta_s,a_s,\lambda)\right)\right] \\ &=\mathbb{E}\left[J(\delta_0,a_0,\lambda)+\int_0^T e^{-rs}\left(dC_s^a-rJ(\delta_s,a_s,\lambda)\right.\right. \\ &+\left(J(\delta_s,a_s,\lambda)-J(\delta_{s-},a_s,\lambda)\right)dN_s^\alpha \\ &\left.\left.+\left(J(\delta_s,a_s+q_s\left[\left(\delta_{s-},a_{s-},\lambda\right),\left(\delta_s',a_s',\lambda'\right)\right],\lambda)-J(\delta_s,a_s,\lambda)\right)dN_s^\lambda\right)\right],\end{aligned}$$

where N_s^α and N_s^λ are counting processes that govern the arrivals of idiosyncratic shocks and of potential counterparties, respectively. Note that any transfer of the numéraire at an arrival time of N^λ is reflected by C^a according to (B.8).

The next step is to calculate the stochastic integrals containing the counting processes. The condition (B.4b) implies that,

$$\int_0^T |J(\delta_s, a_s, \lambda) - J(\delta_{s-}, a_s, \lambda)|\, ds \leq \sup_{s,s' \in [0,T]} |J(\delta_{s'}, a_{s'}, \lambda) - J(\delta_s, a_s, \lambda)|\, T < \infty.$$

Corollary $C4$ of Brémaud (1981, p. 235), in turn, implies that

$$\mathbb{E}\left[\int_0^T e^{-rs}\left(J(\delta_s, a_s, \lambda) - J(\delta_{s-}, a_s, \lambda)\right) dN_s^\alpha\right] = \mathbb{E}\left[\int_0^T e^{-rs}\alpha\left\{\int_{\delta_L}^{\delta_H}\left(J(\delta_s, a_s, \lambda) - J(\delta_{s-}, a_s, \lambda)\right) f(\delta_s')d\delta_s'\right\} ds\right].$$

Similarly,

$$\begin{aligned}&\mathbb{E}\left[\int_0^T e^{-rs}(J(\delta_s, a_s + q_s\left[(\delta_{s-}, a_{s-}, \lambda),(\delta_s', a_s', \lambda')\right], \lambda) - J(\delta_s, a_s, \lambda))\, dN_s^\lambda\right]\\ &= \mathbb{E}\left[\int_0^T e^{-rs}\left\{\int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} m\,(\lambda, \lambda_s')(J(\delta_s, a_s + q_s[(\delta_{s-}, a_{s-}, \lambda),(\delta_s', a_s', \lambda')], \lambda)\right.\right.\\ &\qquad\left.\left. - J(\delta_s, a_s, \lambda))\, \Phi(d\delta_s', da_s', d\lambda_s')\right\} ds\right].\end{aligned}$$

Using these equalities in (B.9),

$$\begin{aligned}&\mathbb{E}\left[\int_0^T e^{-rs}dC_s^a + e^{-rT}J(\delta_T, a_T, \lambda)\right] = \mathbb{E}\left[J(\delta_0, a_0, \lambda) + \int_0^T e^{-rs}dC_s^a\right.\\ &+ \int_0^T e^{-rs}\left(\alpha\int_{\delta_L}^{\delta_H}\left(J(\delta_s, a_s, \lambda) - J(\delta_{s-}, a_s, \lambda)\right) f(\delta_s')d\delta_s' - rJ(\delta_s, a_s, \lambda)\right.\\ &+ \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} m\,(\lambda, \lambda_s')\left(J(\delta_s, a_s + q_s\left[(\delta_{s-}, a_{s-}, \lambda), (\delta_s', a_s', \lambda')\right], \lambda)\right.\\ &\left.\left. - J(\delta_s, a_s, \lambda)\right) \Phi(d\delta_s', da_s', d\lambda_s')\right) ds\Big]\end{aligned}$$

$$\leq \mathbb{E}\left[J(\delta_0, a_0, \lambda) + \sup_{C} \left\{ \int_0^T e^{-rs} dC_s \right.\right.$$
$$+ \int_0^T e^{-rs} \left(\alpha \int_{\delta_L}^{\delta_H} \left(J(\delta_s, a_s, \lambda) - J(\delta_{s-}, a_s, \lambda)\right) f(\delta_s') d\delta_s' - rJ(\delta_s, a_s, \lambda) \right.$$
$$+ \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} m\left(\lambda, \lambda_s'\right) \left(J\left(\delta_s, a_s + q_s\left[\left(\delta_{s-}, a_{s-}, \lambda\right), \left(\delta_s', a_s', \lambda'\right)\right], \lambda\right)\right.$$
$$\left.\left.\left. - J(\delta_s, a_s, \lambda)\right) \Phi(d\delta_s', da_s', d\lambda_s')\right) ds \right\}\right] = J(\delta_0, a_0, \lambda).$$

This means that, at any future meeting date τ^n, $n \in \mathbb{N}$,

$$J(\delta_0, a_0, \lambda) \geq \mathbb{E}\left[\int_0^{\tau^n} e^{-rt} dC_t^a \right] + \mathbb{E}\left[e^{-r\tau^n} J(\delta_{\tau^n}, a_{\tau^n}, \lambda) \right].$$

Then, letting $n \to \infty$ and using the transversality condition (B.4a), I find $J(\delta_0, a_0, \lambda) \geq U\left(C^a\right)$. Since $J(\delta_0, a_0, \lambda) = U\left(C^*\right)$, where C^* is the consumption process associated with the candidate equilibrium strategy, I have established optimality.

Existence and Uniqueness

In Appendix A, I construct a solution to the HJB equation (B.7) for $J(\delta, a, \lambda)$. Taking as given the equilibrium joint cdf $\Phi(\delta, a, \lambda)$ of taste types, asset positions, and speed types, here I provide a formal proof to establish that (B.7) does not admit another real solution. The argument does not use standard fixed point tools for dynamic programming unlike the earlier models with unrestricted asset positions, such as Gârleanu (2009) and Lagos and Rocheteau (2009), because the return function, u, is unbounded below in the support of the equilibrium asset distribution, which is the entire real line. Thus, I prove the uniqueness without relying on the boundedness of the return function along feasible paths. In particular, I follow the metric approach of Rincón-Zapatero and Rodríguez-Palmero (2003) and Martins-da Rocha and Vailakis (2010) and show that the operator defined by (B.7) is a contraction. Then, the contraction mapping theorem implies that it has a unique fixed point.

Let $\omega_- \leq \omega_+ < \omega$ be three real-valued continuous functions defined on $\mathcal{T}$, such that

$$\omega_-\left(\delta, a, \lambda\right) = \frac{1}{r} \frac{ru\left(\delta, a\right) + \alpha u\left(\overline{\delta}, a\right)}{r + \alpha},$$
$$\omega_+\left(\delta, a, \lambda\right) = \frac{1}{r} \frac{ru\left(\delta, a\right) + \left(\alpha + \frac{1}{2}m\left(\lambda, \Lambda\right)\right) u\left(\delta_H, \frac{\delta_H}{\kappa}\right)}{r + \alpha + \frac{1}{2}m\left(\lambda, \Lambda\right)},$$

and

$$\omega(\delta, a, \lambda) = \frac{\eta}{r} u\left(\delta_H, \frac{\delta_H}{\kappa}\right),$$

where $\eta > 1$ is an arbitrarily large real number. Now consider the following metric on $C(\mathcal{T}) \equiv \{f : \mathcal{T} \to \mathbb{R} \,| f \text{ is continuous}, \omega_- \leq f \leq \omega_+\}$:

$$d(f, g) = \sup_{x \in \mathcal{T}} \left| \log\left(\frac{f - \omega}{\omega_+ - \omega}(x)\right) - \log\left(\frac{g - \omega}{\omega_+ - \omega}(x)\right) \right|, \quad f, g \in C(\mathcal{T}).$$

Here, ω_- and ω_+ functions provide natural lower and upper bounds for the candidate equilibrium value function, respectively. The function ω_- is calculated assuming the investor cannot trade, i.e., $m(\lambda, \lambda') = 0$. The function ω (which is in fact a constant) is the value of receiving the highest possible utility flow forever, scaled up by the coefficient η. As will be clear shortly, choosing η large enough makes the operator, defined by (B.7), a contraction that maps $C(\mathcal{T})$ into itself. Lastly, ω_+ is the value of of receiving the highest possible utility flow forever after the operator is applied once.

LEMMA 5 $(C(\mathcal{T}), d)$ *is a complete metric space.*

PROOF: That d is a metric is obvious (Stokey and Lucas, 1989, p. 44). It suffices to show that every Cauchy sequence in $C(\mathcal{T})$ converges to a function in $C(\mathcal{T})$. I do it in three steps. First, given a Cauchy sequence $\{f_n\}$ in $C(\mathcal{T})$ I find a candidate limiting function f. Second, I show that $\{f_n\}$ converges to the candidate functions in d metric. Finally, I show that the candidate function f belongs to $C(\mathcal{T})$.

1. Given a Cauchy sequence $\{f_n\}$ in $C(\mathcal{T})$, it holds that $d(f_n, f_m) \to 0$ as $n, m \to \infty$. Fix $x \in \mathcal{T}$; then the sequence of real numbers $\left\{\log\left(\frac{f_n - \omega}{\omega_+ - \omega}(x)\right)\right\}$ satisfies the Cauchy criterion; and by the completeness of real numbers, it converges to a limit point—call it $g(x)$. The limiting values define a function $g : \mathcal{T} \to \mathbb{R}_+$. I take our "candidate" limiting function for $\{f_n\}$ to be $f \equiv e^g(\omega_+ - \omega) + \omega$.
2. $$d(f_n, f) = d(f_n, e^g(\omega_+ - \omega) + \omega) = \sup_{x \in \mathcal{T}} \left| \log\left(\frac{f - \omega}{\omega_+ - \omega}(x)\right) - g(x) \right|$$
   tends to zero as $n \to \infty$.
3. I have to show that f is continuous and $\omega_- \leq f \leq \omega_+$. To prove that f is continuous, I must show that for every $\varepsilon > 0$ and every $x \in \mathcal{T}$, there exists $\zeta > 0$ such that
   $$\left| \log\left(\frac{f - \omega}{\omega_+ - \omega}(x)\right) - \log\left(\frac{f - \omega}{\omega_+ - \omega}(y)\right) \right| < \varepsilon \text{ if } \|x - y\|_E < \zeta,$$

where $\|.\|_E$ is the Euclidean sup norm on $\mathbb{R}^3$. Let ε and x be given. Choose k so that $d(f, f_k) < \varepsilon/3$; since $f_n \to f$, such a choice possible. Then choose ζ so that

$$\|x - y\|_E < \zeta \text{ implies } \left| \log\left(\frac{f_k - \omega}{\omega_+ - \omega}(x)\right) - \log\left(\frac{f_k - \omega}{\omega_+ - \omega}(y)\right) \right| < \varepsilon/3.$$

Since f_k is continuous, such a choice is possible. Then,

$$\begin{aligned} &\left| \log\left(\frac{f - \omega}{\omega_+ - \omega}(x)\right) - \log\left(\frac{f - \omega}{\omega_+ - \omega}(y)\right) \right| \\ &\leq \left| \log\left(\frac{f - \omega}{\omega_+ - \omega}(x)\right) - \log\left(\frac{f_k - \omega}{\omega_+ - \omega}(x)\right) \right| \\ &+ \left| \log\left(\frac{f_k - \omega}{\omega_+ - \omega}(x)\right) - \log\left(\frac{f_k - \omega}{\omega_+ - \omega}(y)\right) \right| \\ &+ \left| \log\left(\frac{f_k - \omega}{\omega_+ - \omega}(y)\right) - \log\left(\frac{f - \omega}{\omega_+ - \omega}(y)\right) \right| \\ &\leq 2d(f, f_k) + \left| \log\left(\frac{f_k - \omega}{\omega_+ - \omega}(x)\right) - \log\left(\frac{f_k - \omega}{\omega_+ - \omega}(y)\right) \right| < \varepsilon, \end{aligned}$$

which implies that f is continuous. Lastly, $\omega_- \leq f \leq \omega_+$ follows from the fact that $\leq$ is a continuous relation.

Q.E.D.

LEMMA 6 *Suppose $J \in C(\mathcal{T})$. Let $S : \mathcal{T}^2 \times \mathbb{R} \to \mathbb{R}$ and $S^* : \mathcal{T}^2 \to \mathbb{R}$ be*

$$S\left[(\delta, a, \lambda), (\delta', a', \lambda'), q\right] = J(\delta, a + q, \lambda) + J(\delta', a' - q, \lambda')$$

and

$$S^*\left[(\delta, a, \lambda), (\delta', a', \lambda')\right] = \max_{q \in \mathbb{R}} S\left[(\delta, a, \lambda), (\delta', a', \lambda'), q\right],$$

respectively. Then, S^* is continuous.

PROOF: This result is a direct application of Theorem 3.1 of Montes-de Oca and Lemus-Rodríguez (2012), which is a generalization of Berge's theorem that permits to deal with optimization problems with unbounded objective function and noncompact restrictions set. The theorem states that if the restrictions set is a closed-valued and continuous correspondence and if the objective function of the *minimization* problem is continuous, inf-compact, and satisfies the Moment Condition (MC), then the minimized values define a continuous function on the parameter space (Montes-de Oca and Lemus-Rodríguez, 2012, p. 272).

Our restrictions correspondence takes the value of $\mathbb{R}$ for all $[(\delta, a, \lambda), (\delta', a', \lambda')] \in \mathcal{T}^2$, which means it is closed-valued and continuous. Then, when I show that

$-S\left[(\delta,a,\lambda),(\delta',a',\lambda'),q\right]$ is continuous, satisfies the MC, and is inf-compact, the proof will be complete. Continuity follows from the assumption that $J \in C(\mathcal{T})$. To see the MC holds, let $\mathcal{T}_n = [\delta_L,\delta_H] \times [-n,n] \times [0,M]$ and consider the sequence of compact sets be $\{\mathbb{K}_n\} = \left\{\mathcal{T}_n^2 \times [-n,n]\right\}$ such that $\mathbb{K}_n \uparrow \mathcal{T}^2 \times \mathbb{R}$. Since $J \in C(\mathcal{T})$,

$$\begin{aligned} -\omega_+(\delta,a+q,\lambda) - \omega_+(\delta',a'-q,\lambda') &\leq -S\left[(\delta,a,\lambda),(\delta',a',\lambda'),q\right] \\ &\leq -\omega_-(\delta,a+q,\lambda) - \omega_-(\delta',a'-q,\lambda'). \end{aligned}$$

Thus,

$$\begin{aligned} &\inf_{[(\delta,a,\lambda),(\delta',a',\lambda'),q]\notin\mathbb{K}_n} -\omega_+(\delta,a+q,\lambda) - \omega_+(\delta',a'-q,\lambda') \\ &\leq \inf_{[(\delta,a,\lambda),(\delta',a',\lambda'),q]\notin\mathbb{K}_n} -S\left[(\delta,a,\lambda),(\delta',a',\lambda'),q\right] \\ &\leq \inf_{[(\delta,a,\lambda),(\delta',a',\lambda'),q]\notin\mathbb{K}_n} -\omega_-(\delta,a+q,\lambda) - \omega_-(\delta',a'-q,\lambda'). \end{aligned}$$

Since

$$\lim_{n\to\infty}\left(\inf_{[(\delta,a,\lambda),(\delta',a',\lambda'),q]\notin\mathbb{K}_n} -\omega_+(\delta,a+q,\lambda) - \omega_+(\delta',a'-q,\lambda')\right) = \infty$$

and

$$\lim_{n\to\infty}\left(\inf_{[(\delta,a,\lambda),(\delta',a',\lambda'),q]\notin\mathbb{K}_n} -\omega_-(\delta,a+q,\lambda) - \omega_-(\delta',a'-q,\lambda')\right) = \infty,$$

then, by Squeeze Theorem,

$$\lim_{n\to\infty}\left(\inf_{[(\delta,a,\lambda),(\delta',a',\lambda'),q]\notin\mathbb{K}_n} -S\left[(\delta,a,\lambda),(\delta',a',\lambda'),q\right]\right) = \infty,$$

which means the MC holds. Finally, to show that $-S$ is inf-compact, I must show that the set $\{q \in \mathbb{R} \mid S\left[(\delta,a,\lambda),(\delta',a',\lambda'),q\right] \geq c\}$ is compact for every $[(\delta,a,\lambda),(\delta',a',\lambda')] \in \mathcal{T}^2$ and $c \in \mathbb{R}$ (Montes-de Oca and Lemus-Rodríguez, 2012, p. 271). It suffices to show that the set is bounded and closed. Note that $J \in C(\mathcal{T})$ implies

$$\begin{aligned} &\{q \in \mathbb{R} \mid S\left[(\delta,a,\lambda),(\delta',a',\lambda'),q\right] \geq c\} \\ &\qquad\qquad \subset \{q \in \mathbb{R} \mid \omega_+(\delta,a+q,\lambda) + \omega_+(\delta',a'-q,\lambda') \geq c\}. \end{aligned}$$

As $\omega_+(\delta,a+q,\lambda) + \omega_+(\delta',a'-q,\lambda')$ is a concave-quadratic function of q, the latter set is bounded, which implies that the former is bounded as well. To see closedness, let $\{q_n\} \subset \{q \in \mathbb{R} \mid S\left[(\delta,a,\lambda),(\delta',a',\lambda'),q\right] \geq c\}$ and $q_n \to q^*$. Then, because $S\left[(\delta,a,\lambda),(\delta',a',\lambda'),q_n\right] \geq c$ and the weak inequality is preserved under the limit, we have $S\left[(\delta,a,\lambda),(\delta',a',\lambda'),q^*\right] \geq c$, which completes the proof. *Q.E.D.*

LEMMA 7 *Suppose Φ is a joint cdf such that*

$$\int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} f(x)\, d\Phi(x) < \infty \tag{B.10}$$

for any $f \in C(\mathcal{T})$. Then, in the set $C(\mathcal{T})$, there exists a unique solution to (3.13) (or B.7).

PROOF: Rewrite (3.13) as

$$J(\delta,a,\lambda) = \frac{1}{r+\alpha+\frac{1}{2}m(\lambda,\Lambda)} \Bigg(u(\delta,a) + \alpha \int_{\delta_L}^{\delta_H} J(\delta',a,\lambda) f(\delta') d\delta' + \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{1}{2} m(\lambda,\lambda') \Big[\max_q \{J(\delta,a+q,\lambda) + J(\delta',a'-q,\lambda') - J(\delta',a',\lambda')\}\Big] \Phi(d\delta',da',d\lambda') \Bigg). \tag{B.11}$$

The RHS of (B.11) defines a mapping O:

$$(OJ)(\delta,a,\lambda) = \frac{1}{r+\alpha+\frac{1}{2}m(\lambda,\Lambda)} \Bigg(u(\delta,a) + \alpha \int_{\delta_L}^{\delta_H} J(\delta',a,\lambda) f(\delta') d\delta' + \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{1}{2} m(\lambda,\lambda') \Big[\max_q \{J(\delta,a+q,\lambda) + J(\delta',a'-q,\lambda') - J(\delta',a',\lambda')\}\Big] \Phi(d\delta',da',d\lambda') \Bigg).$$

I want to show that there exists a unique solution J to $OJ = J$. I do so in two steps. In the first step, I establish that $O : C(\mathcal{T}) \to C(\mathcal{T})$. In the second step, I show that O is a contraction. Then, it follows from the contraction mapping theorem that O has a unique fixed point $J \in C(\mathcal{T})$.

1. Suppose $J \in C(\mathcal{T})$, then by Lemma 6 OJ is continuous. Next, I show $\omega_- \le OJ \le \omega_+$. For $\omega_- \le OJ$, it suffices to show that

$$u(\delta,a) + \alpha \int_{\delta_L}^{\delta_H} J(\delta',a,\lambda) f(\delta') d\delta' - (r+\alpha)\,\omega_-(\delta,a,\lambda) + \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{1}{2} m(\lambda,\lambda') \Big[\max_q \{J(\delta,a+q,\lambda) + J(\delta',a'-q,\lambda') - \omega_-(\delta,a,\lambda) - J(\delta',a',\lambda')\}\Big] \Phi(d\delta',da',d\lambda') \ge 0$$

for all $(\delta, a, \lambda) \in \mathcal{T}$. The expression in the last two lines is weakly larger than zero by the choice of $q = 0$. And,

$$
\begin{aligned}
& u(\delta, a) + \alpha \int_{\delta_L}^{\delta_H} J(\delta', a, \lambda) f(\delta') d\delta' - (r+\alpha)\, \omega_-(\delta, a, \lambda) \\
& \geq u(\delta, a) + \alpha \int_{\delta_L}^{\delta_H} \omega_-(\delta', a, \lambda) f(\delta') d\delta' - (r+\alpha)\, \omega_-(\delta, a, \lambda) = 0,
\end{aligned}
$$

which implies $\omega_- \leq OJ$. For $OJ \leq \omega_+$, it suffices to show that

$$
\begin{aligned}
& u(\delta, a) - u(\delta, a) + \alpha \int_{\delta_L}^{\delta_H} J(\delta', a, \lambda) f(\delta') d\delta' - \alpha \frac{\eta}{r} u\left(\delta_H, \frac{\delta_H}{\kappa}\right) \\
& + \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{1}{2} m(\lambda, \lambda') \Bigg[\max_q \ \Big\{ J(\delta, a+q, \lambda) \ + J(\delta', a'-q, \lambda') \\
& \qquad - \frac{\eta}{r} u\left(\delta_H, \frac{\delta_H}{\kappa}\right) - J(\delta', a', \lambda') \Big\} \Bigg] \Phi(d\delta', da', d\lambda') \leq 0
\end{aligned}
$$

for all $(\delta, a, \lambda) \in \mathcal{T}$. That the first line is weakly smaller than zero is obvious. And,

$$
\begin{aligned}
& \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{1}{2} m(\lambda, \lambda') \Bigg[\max_q \ \Big\{ J(\delta, a+q, \lambda) \ + J(\delta', a'-q, \lambda') \\
& \quad - \frac{\eta}{r} u\left(\delta_H, \frac{\delta_H}{\kappa}\right) - J(\delta', a', \lambda') \Big\} \Bigg] \Phi(d\delta', da', d\lambda') \\
& \leq \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{1}{2} m(\lambda, \lambda') \Bigg[\max_q \ \Big\{ \omega_+(\delta, a+q, \lambda) \ + \omega_+(\delta', a'-q, \lambda') \\
& \quad - \frac{\eta}{r} u\left(\delta_H, \frac{\delta_H}{\kappa}\right) - \omega_-(\delta', a', \lambda') \Big\} \Bigg] \Phi(d\delta', da', d\lambda')
\end{aligned}
$$

Thanks to (B.10), it is possible to choose η sufficiently large so that the RHS of the previous inequality is weakly smaller than zero, which implies $OJ \leq \omega_+$. Hence, $O : C(\mathcal{T}) \to C(\mathcal{T})$.

2. I next show that O is a contraction mapping. The main property of the mapping O, I use in this proof is *convexity*, i.e., for $\zeta \in [0, 1]$,

$$
O\left(\zeta J^A + (1-\zeta) J^B\right) \leq \zeta O J^A + (1-\zeta) O J^B,
$$

thanks to the convexity of the max operator. Moreover, I make use of a version of *monotonicity*. It is easy to see that O does not have to be monotone because of the "$-J(\delta', a', \lambda')$" term at the end of its definition. As a result, I use the monotonicity of

$$(Of)(\delta, a, \lambda) + \frac{\int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{1}{2} m(\lambda, \lambda') f(\delta', a', \lambda') \Phi(d\delta', da', d\lambda')}{r + \alpha + \frac{1}{2} m(\lambda, \Lambda)},$$

which is the property that $J^A \leq J^B$ implies

$$\begin{aligned}
&\left(OJ^A\right)(\delta, a, \lambda) + \frac{\int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{1}{2} m(\lambda, \lambda') J^A(\delta', a', \lambda') \Phi(d\delta', da', d\lambda')}{r + \alpha + \frac{1}{2} m(\lambda, \Lambda)} \\
&\leq \left(OJ^B\right)(\delta, a, \lambda) + \frac{\int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{1}{2} m(\lambda, \lambda') J^B(\delta', a', \lambda') \Phi(d\delta', da', d\lambda')}{r + \alpha + \frac{1}{2} m(\lambda, \Lambda)}.
\end{aligned}$$

Following the same steps as in Rincón-Zapatero and Rodríguez-Palmero (2003, p. 1553),

$$J^A \leq e^{-d\left(J^A, J^B\right)} J^B + \left(1 - e^{-d\left(J^A, J^B\right)}\right)\omega \tag{B.12}$$

for all J^A, J^B with $d\left(J^A, J^B\right)$ being well defined. Using the monotonicity and convexity,

$$\begin{aligned}
&\left(OJ^A\right)(\delta, a, \lambda) + \frac{\int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{1}{2} m(\lambda, \lambda') J^A(\delta', a', \lambda') \Phi(d\delta', da', d\lambda')}{r + \alpha + \frac{1}{2} m(\lambda, \Lambda)} \\
&\leq O\left(e^{-d\left(J^A, J^B\right)} J^B + \left(1 - e^{-d\left(J^A, J^B\right)}\right)\omega\right)(\delta, a, \lambda) \\
&+ \frac{\left(1 - e^{-d\left(J^A, J^B\right)}\right) \frac{1}{2} m(\lambda, \Lambda)}{r + \alpha + \frac{1}{2} m(\lambda, \Lambda)}\omega \\
&+ \frac{e^{-d\left(J^A, J^B\right)} \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{1}{2} m(\lambda, \lambda') J^B(\delta', a', \lambda') \Phi(d\delta', da', d\lambda')}{r + \alpha + \frac{1}{2} m(\lambda, \Lambda)}
\end{aligned}$$

$$\leq e^{-d\left(J^A,J^B\right)} OJ^B(\delta,a,\lambda) + \left(1 - e^{-d\left(J^A,J^B\right)}\right) O\omega(\delta,a,\lambda)$$
$$+ \frac{\left(1 - e^{-d\left(J^A,J^B\right)}\right) \frac{1}{2} m\left(\lambda,\Lambda\right)}{r + \alpha + \frac{1}{2} m\left(\lambda,\Lambda\right)} \omega$$
$$+ \frac{e^{-d\left(J^A,J^B\right)} \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{1}{2} m\left(\lambda,\lambda'\right) J^B(\delta',a',\lambda') \Phi(d\delta',da',d\lambda')}{r + \alpha + \frac{1}{2} m\left(\lambda,\Lambda\right)}.$$

Defining

$$h(\delta,a,\lambda|J^A,J^B) \equiv$$
$$\frac{\int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{m(\lambda,\lambda')}{m(\lambda,\Lambda)} \left[e^{-d\left(J^A,J^B\right)} J^B(\delta',a',\lambda') - J^A(\delta',a',\lambda')\right] \Phi(d\delta',da',d\lambda')}{e^{-d(J^A,J^B)} - 1}$$

and suppressing (δ,a,λ)s,

$$OJ^A \leq e^{-d\left(J^A,J^B\right)} OJ^B + \left(1 - e^{-d\left(J^A,J^B\right)}\right) O\omega$$
$$+ \left(1 - e^{-d\left(J^A,J^B\right)}\right) \frac{\frac{1}{2} m\left(\lambda,\Lambda\right)}{r + \alpha + \frac{1}{2} m\left(\lambda,\Lambda\right)} \omega$$
$$+ \left(e^{-d\left(J^A,J^B\right)} - 1\right) \frac{\frac{1}{2} m\left(\lambda,\Lambda\right)}{r + \alpha + \frac{1}{2} m\left(\lambda,\Lambda\right)} h.$$

Applying (B.12) to the second and the last terms on the RHS,[2]

$$OJ^A \leq e^{-d\left(J^A,J^B\right)} OJ^B$$
$$+ \left(1 - e^{-d\left(J^A,J^B\right)}\right) \left[e^{-d\left(OJ^B,O\omega\right)} OJ^B + \left(1 - e^{-d\left(OJ^B,O\omega\right)}\right) \omega\right]$$
$$+ \left(1 - e^{-d\left(J^A,J^B\right)}\right) \frac{\frac{1}{2} m\left(\lambda,\Lambda\right)}{r + \alpha + \frac{1}{2} m\left(\lambda,\Lambda\right)} \omega$$
$$+ \left(e^{-d\left(J^A,J^B\right)} - 1\right) \frac{\frac{1}{2} m\left(\lambda,\Lambda\right)}{r + \alpha + \frac{1}{2} m\left(\lambda,\Lambda\right)}$$
$$\left[e^{-d\left(OJ^B,h\right)} OJ^B + \left(1 - e^{-d\left(OJ^B,h\right)}\right) \omega\right].$$

[2] Note that h does not have to be an element of the set $C\left(\mathcal{T}\right)$ to write down these intermediate inequalities. The only requirement we need is $h < \omega$, which follows from $J^A, J^B < \omega$, (B.10), and the fact that η is arbitrarily large.

Subtracting ω from both sides and dividing both sides by $\omega_+ - \omega$,

$$\begin{aligned}
&\frac{OJ^A - \omega}{\omega_+ - \omega} \geq \frac{OJ^B - \omega}{\omega_+ - \omega} \\
&\times \Bigg[e^{-d\left(OJ^B, O\omega\right)} - \frac{\frac{1}{2}m\left(\lambda, \Lambda\right)}{r + \alpha + \frac{1}{2}m\left(\lambda, \Lambda\right)} e^{-d\left(OJ^B, h\right)} \\
&+ \left(1 - e^{-d\left(OJ^B, O\omega\right)}\right) \\
&\qquad \left(1 - e^{-d\left(OJ^B, O\omega\right)} + \frac{\frac{1}{2}m\left(\lambda, \Lambda\right)}{r + \alpha + \frac{1}{2}m\left(\lambda, \Lambda\right)} e^{-d\left(OJ^B, h\right)}\right)\Bigg].
\end{aligned}$$

Taking the logarithm of both sides,

$$\log\left(\frac{OJ^A - \omega}{\omega_+ - \omega}\right) \geq \log\left(z\right) + \log\left(\frac{OJ^B - \omega}{\omega_+ - \omega}\right), \tag{B.13}$$

where

$$\begin{aligned}
&z = e^{-d\left(OJ^B, O\omega\right)} - \frac{\frac{1}{2}m\left(\lambda, \Lambda\right)}{r + \alpha + \frac{1}{2}m\left(\lambda, \Lambda\right)} e^{-d\left(OJ^B, h\right)} \\
&+ \left(1 - e^{-d\left(OJ^B, O\omega\right)}\right) \\
&\qquad \left(1 - e^{-d\left(OJ^B, O\omega\right)} + \frac{\frac{1}{2}m\left(\lambda, \Lambda\right)}{r + \alpha + \frac{1}{2}m\left(\lambda, \Lambda\right)} e^{-d\left(OJ^B, h\right)}\right).
\end{aligned}$$

Appendix B of Rincón-Zapatero and Rodríguez-Palmero (2003, p. 1554) shows that $\log\left(a + e^{-x}(1-a)\right) \geq -\left(1-a\right)x$. Applying to (B.13),

$$\begin{aligned}
&\log\left(\frac{OJ^B - \omega}{\omega_+ - \omega}\right) \leq \log\left(\frac{OJ^A - \omega}{\omega_+ - \omega}\right) \\
&+\left(1 - e^{-d\left(OJ^B, O\omega\right)} + \frac{\frac{1}{2}m\left(\lambda, \Lambda\right)}{r + \alpha + \frac{1}{2}m\left(\lambda, \Lambda\right)} e^{-d\left(OJ^B, h\right)}\right) d\left(J^A, J^B\right).
\end{aligned}$$

Applying the same procedure in reverse establishes

$$\begin{aligned}
&\log\left(\frac{OJ^A - \omega}{\omega_+ - \omega}\right) \leq \log\left(\frac{OJ^B - \omega}{\omega_+ - \omega}\right) \\
&+\left(1 - e^{-d\left(OJ^A, O\omega\right)} + \frac{\frac{1}{2}m\left(\lambda, \Lambda\right)}{r + \alpha + \frac{1}{2}m\left(\lambda, \Lambda\right)} e^{-d\left(OJ^A, h'\right)}\right) d\left(J^A, J^B\right),
\end{aligned}$$

where

$$h'(\delta, a, \lambda | J^A, J^B) \equiv$$
$$\frac{\int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \frac{m(\lambda,\lambda')}{m(\lambda,\Lambda)} \left[e^{-d\left(J^A, J^B\right)} J^A(\delta', a', \lambda') - J^B(\delta', a', \lambda') \right] \Phi(d\delta', da', d\lambda')}{e^{-d(J^A,J^B)} - 1}.$$

Thus, to show that O is a contraction, I have to establish that

$$\text{(B.14)} \qquad \sup_{J^A, J^B \in C(\mathcal{T})} \max \left\{ 1 - e^{-d\left(OJ^B, O\omega\right)} + \frac{\frac{1}{2} m(M,\Lambda)}{r + \alpha + \frac{1}{2} m(M,\Lambda)} e^{-d\left(OJ^B, h\right)}, \right.$$
$$\left. 1 - e^{-d\left(OJ^A, O\omega\right)} + \frac{\frac{1}{2} m(M,\Lambda)}{r + \alpha + \frac{1}{2} m(M,\Lambda)} e^{-d\left(OJ^A, h'\right)} \right\} < 1.$$

One can easily verify that

$$\lim_{\eta \to \infty} d(J, O\omega) = \log \left(\frac{r + \alpha + \frac{1}{2} m(M,\Lambda)}{r} \right),$$

$$\lim_{\eta \to \infty} d\left(J, h\left(J^A, J^B\right)\right) = \infty,$$

and

$$\lim_{\eta \to \infty} d\left(J, h'\left(J^A, J^B\right)\right) = \infty$$

for any $J, J^A, J^B \in C(\mathcal{T})$. Then, as $\eta \to \infty$, the LHS of (B.14) approaches $\frac{\alpha + \frac{1}{2} m(M,\Lambda)}{r + \alpha + \frac{1}{2} m(M,\Lambda)}$. Thus, it is possible to choose η sufficiently large so that the LHS of (B.14) is strictly smaller than 1, which implies that O is a contraction mapping on the complete metric space $(C(\mathcal{T}), d)$, as is proven in Lemma 5. Hence, it follows from the *contraction mapping theorem* that O has a unique fixed point $J \in C(\mathcal{T})$ (Theorem 3.2 of Stokey and Lucas, 1989, p. 50).

Q.E.D.

APPENDIX C: PROOFS OMITTED FROM THE PRINTED VERSION

Proof of Proposition 4

Using Proposition 2,

$$\begin{aligned}
\mathcal{GV}(\theta,\lambda) &= \int_0^M \int_{-\infty}^{\infty} m(\lambda,\lambda') \left| q\left[(\theta,\lambda),(\theta',\lambda')\right]\right| g_{\lambda'}(\theta')\,\psi(\lambda')\,d\theta' d\lambda' \\
&= \int_0^M \int_{-\infty}^{\infty} m(\lambda,\lambda') \left| \frac{\widetilde{r}(\lambda')\,\theta - \widetilde{r}(\lambda)\,\theta'}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}\right| g_{\lambda'}(\theta')\,\psi(\lambda')\,d\theta' d\lambda' \\
&= \int_0^M m(\lambda,\lambda') \left\{ \int_{-\infty}^{\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta} \frac{\widetilde{r}(\lambda')\,\theta - \widetilde{r}(\lambda)\,\theta'}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')} g_{\lambda'}(\theta')\,d\theta' \right. \\
&\qquad \left. + \int_{\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta}^{\infty} \frac{\widetilde{r}(\lambda)\,\theta' - \widetilde{r}(\lambda')\,\theta}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')} g_{\lambda'}(\theta')\,d\theta' \right\} \psi(\lambda')\,d\lambda'.
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{NV}(\theta,\lambda) &= \left| \int_0^M \int_{-\infty}^{\infty} m(\lambda,\lambda')\, q\left[(\theta,\lambda),(\theta',\lambda')\right] g_{\lambda'}(\theta')\,\psi(\lambda')\,d\theta' d\lambda' \right| \\
&= \left| \int_0^M \int_{-\infty}^{\infty} m(\lambda,\lambda') \frac{\widetilde{r}(\lambda')\,\theta - \widetilde{r}(\lambda)\,\theta'}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')} g_{\lambda'}(\theta')\,\psi(\lambda')\,d\theta' d\lambda' \right| \\
&= \left| \int_0^M \int_{-\infty}^{\infty} m(\lambda,\lambda') \frac{\widetilde{r}(\lambda')\,\theta}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')} g_{\lambda'}(\theta')\,\psi(\lambda')\,d\theta' d\lambda' \right| \\
&= \left| \int_0^M m(\lambda,\lambda') \frac{\widetilde{r}(\lambda')\,\theta}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')} \psi(\lambda')\,d\lambda' \right| = 2\left(\widetilde{r}(\lambda)-r\right)|\theta|.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathcal{IV}(\theta,\lambda) &= \int_0^M m(\lambda,\lambda') \left\{ \int_{-\infty}^{\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta} \frac{\widetilde{r}(\lambda')\left[\theta-|\theta|\right] - \widetilde{r}(\lambda)\,\theta'}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')} g_{\lambda'}(\theta')\,d\theta' \right. \\
&\qquad \left. + \int_{\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta}^{\infty} \frac{\widetilde{r}(\lambda)\,\theta' - \widetilde{r}(\lambda')\left[\theta+|\theta|\right]}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')} g_{\lambda'}(\theta')\,d\theta' \right\} \psi(\lambda')\,d\lambda'.
\end{aligned}$$

To derive *(i)*, one can take derivative with respect to θ applying the Leibniz rule whenever necessary:

$$\begin{aligned}
\frac{\partial \mathcal{GV}(\theta,\lambda)}{\partial\theta} &= \int_0^M m(\lambda,\lambda')\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}\left[2G_{\lambda'}\left(\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta\right)-1\right]\psi(\lambda')\,d\lambda', \\
\frac{\partial \mathcal{NV}(\theta,\lambda)}{\partial\theta} &= 2\left(\widetilde{r}(\lambda)-r\right)\operatorname{sgn}\theta, \\
\frac{\partial \mathcal{IV}(\theta,\lambda)}{\partial\theta} &= \int_0^M m(\lambda,\lambda')\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}\left[2G_{\lambda'}\left(\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta\right)-1-\operatorname{sgn}\theta\right]\psi(\lambda')d\lambda'.
\end{aligned}$$

Since δ is distributed symmetrically, Equation (3.28) implies $\widehat{g}_\lambda(z)=\widehat{g}_\lambda(-z)$, and hence, θ is distributed symmetrically conditional on λ. Then,

$$\frac{\partial \mathcal{GV}(\theta,\lambda)}{\partial\theta}\begin{cases} <0 & \text{if } \theta<0 \\ =0 & \text{if } \theta=0 \\ >0 & \text{if } \theta>0\end{cases}$$

and the gross volume is minimized at $\theta=0$. The behavior of the net volume is also the same. However, the intermediation volume behaves oppositely:

$$\frac{\partial \mathcal{IV}(\theta,\lambda)}{\partial\theta}\begin{cases} <0 & \text{if } \theta>0 \\ =0 & \text{if } \theta=0 \\ >0 & \text{if } \theta<0,\end{cases}$$

hence the intermediation volume is maximized at $\theta=0$.

To derive *(ii)*, one takes derivative with respect to λ using Lemma 1 and applying the chain rule and the Leibniz rule whenever necessary:

$$\begin{aligned}
\frac{\partial \mathcal{GV}(\theta,\lambda)}{\partial\lambda} = & \left[\int_0^M m(\lambda,\lambda')\frac{\widetilde{r}(\lambda')}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda')\right)^2}\right. \\
& \left\{\left(\mathbb{E}_g\left[\theta'|\theta'>\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta,\lambda'\right]+\theta\right)\left(1-G_{\lambda'}\left(\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta\right)\right)\right. \\
& \left.\left.-\left(\mathbb{E}_g\left[\theta'|\theta'<\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta,\lambda'\right]+\theta\right)G_{\lambda'}\left(\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta\right)\right\}\psi(\lambda')d\lambda'\right]\widetilde{r}'(\lambda) \\
& +\int_0^M\frac{\partial m(\lambda,\lambda')}{\partial\lambda}\left\{\int_{-\infty}^{\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta}\frac{\widetilde{r}(\lambda')\theta-\widetilde{r}(\lambda)\theta'}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}g_{\lambda'}(\theta')\,d\theta'\right. \\
& \left.+\int_{\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta}^{\infty}\frac{\widetilde{r}(\lambda)\theta'-\widetilde{r}(\lambda')\theta}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}g_{\lambda'}(\theta')\,d\theta'\right\}\psi(\lambda')\,d\lambda',
\end{aligned}$$

$$\frac{\partial \mathcal{NV}(\theta,\lambda)}{\partial \lambda} = 2\widetilde{r}'(\lambda)|\theta|,$$

$$\begin{aligned}
\frac{\partial \mathcal{IV}(\theta,\lambda)}{\partial \lambda} &= \left[\int_0^M m(\lambda,\lambda')\frac{\widetilde{r}(\lambda')}{(\widetilde{r}(\lambda)+\widetilde{r}(\lambda'))^2}\left\{-\theta\left(2G_{\lambda'}\left(\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta\right)-1\right)+|\theta|\right.\right.\\
&\quad+\mathbb{E}_g\left[\theta'|\theta'>\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta,\lambda'\right]\left(1-G_{\lambda'}\left(\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta\right)\right)\\
&\quad\left.\left.-\mathbb{E}_g\left[\theta'|\theta'<\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta,\lambda'\right]\left(G_{\lambda'}\left(\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta\right)\right)\right\}\psi(\lambda')\,d\lambda'\right]\widetilde{r}'(\lambda)\\
&\quad+\int_0^M\frac{\partial m(\lambda,\lambda')}{\partial\lambda}\left\{\int_{-\infty}^{\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta}\frac{\widetilde{r}(\lambda')[\theta-|\theta|]-\widetilde{r}(\lambda)\theta'}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}g_{\lambda'}(\theta')\,d\theta'\right.\\
&\quad\left.+\int_{\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta}^{\infty}\frac{\widetilde{r}(\lambda)\theta'-\widetilde{r}(\lambda')[\theta+|\theta|]}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}g_{\lambda'}(\theta')\,d\theta'\right\}\psi(\lambda')\,d\lambda'.
\end{aligned}$$

Using the symmetry of θ around 0 for all λs, $\mathbb{E}_g\left[\theta'|\theta'>\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta,\lambda'\right]>0$ and $\mathbb{E}_g\left[\theta'|\theta'<\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta,\lambda'\right]<0$. Therefore, the first term of $\frac{\partial \mathcal{GV}(\theta,\lambda)}{\partial\lambda}$ is strictly positive. Since $m(\lambda,\lambda')$ is a linear increasing function of λ, the second term is strictly positive as well, impliying $\frac{\partial \mathcal{GV}(\theta,\lambda)}{\partial\lambda}>0$. $\frac{\partial \mathcal{NV}(\theta,\lambda)}{\partial\lambda}\geq 0$ (with equality if $\theta=0$) by the definition of absolute value. $-\theta\left(2G_{\lambda'}\left(\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta\right)-1\right)+|\theta|\geq 0$ by the definition of absolute value and the symmetry of θ around 0. The second line of $\frac{\partial \mathcal{IV}(\theta,\lambda)}{\partial\lambda}$ is strictly positive by the same argument that is used for the first term of $\frac{\partial \mathcal{GV}(\theta,\lambda)}{\partial\lambda}$, implying the first term of $\frac{\partial \mathcal{IV}(\theta,\lambda)}{\partial\lambda}$ (sum of first two lines) is strictly positive. Since $m(\lambda,\lambda')$ is a linear increasing function of λ, the second term is weakly positive, implying $\frac{\partial \mathcal{IV}(\theta,\lambda)}{\partial\lambda}>0$.

Finally, to derive *(iii)*, one takes derivative with respect to λ using Lemma 1 and applying the chain rule and the Leibniz rule whenever necessary:

$$\begin{aligned}
\frac{\partial \mathcal{GV}^{pm}(\theta,\lambda)}{\partial\lambda} &= \widetilde{r}'(\lambda)\left[\int_0^M\frac{m(\lambda,\lambda')}{m(\lambda,\Lambda)}\frac{\widetilde{r}(\lambda')}{(\widetilde{r}(\lambda)+\widetilde{r}(\lambda'))^2}\right.\\
&\quad\left\{\left(\mathbb{E}_g\left[\theta'|\theta'>\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta,\lambda'\right]+\theta\right)\left(1-G_{\lambda'}\left(\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta\right)\right)\right.\\
&\quad\left.\left.-\left(\mathbb{E}_g\left[\theta'|\theta'<\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta,\lambda'\right]+\theta\right)G_{\lambda'}\left(\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)}\theta\right)\right\}\psi(\lambda')d\lambda'\right],
\end{aligned}$$

$$\frac{\partial \mathcal{NV}^{pm}(\theta,\lambda)}{\partial\lambda} = \frac{2|\theta|}{(m(\lambda,\Lambda))^2\left(1+\int_0^M \frac{1}{2}m(\lambda,\lambda')\frac{\widetilde{r}(\lambda')}{(\widetilde{r}(\lambda)+\widetilde{r}(\lambda'))^2}\psi(\lambda')d\lambda'\right)}$$

$$\left\{ m(\lambda,\Lambda)\int_0^M \frac{1}{2}\frac{\partial m(\lambda,\lambda')}{\partial\lambda}\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}\psi(\lambda')d\lambda' \right.$$

$$-\frac{\partial m(\lambda,\Lambda)}{\partial\lambda}\int_0^M \frac{1}{2}m(\lambda,\lambda')\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}\psi(\lambda')d\lambda'$$

$$-\frac{\partial m(\lambda,\Lambda)}{\partial\lambda}\int_0^M \frac{1}{2}m(\lambda,\lambda')\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}\psi(\lambda')d\lambda'$$

$$\left.\int_0^M \frac{1}{2}m(\lambda,\lambda')\frac{\widetilde{r}(\lambda')}{(\widetilde{r}(\lambda)+\widetilde{r}(\lambda'))^2}\psi(\lambda')d\lambda'\right\}.$$

Strict positivity of $\frac{\partial\mathcal{GV}^{pm}(\theta,\lambda)}{\partial\lambda}$ follows from the strict positivity of the first term of $\frac{\partial\mathcal{GV}(\theta,\lambda)}{\partial\lambda}$. Since $m(\lambda,\lambda') = 2\lambda\frac{\lambda'}{\Lambda}$, then $\frac{\partial\mathcal{NV}^{pm}(\theta,\lambda)}{\partial\lambda} \le 0$ (with equality if $\theta = 0$). The strict positivity of $\frac{\partial\mathcal{IV}^{pm}(\theta,\lambda)}{\partial\lambda}$ follows from $\frac{\partial\mathcal{GV}^{pm}(\theta,\lambda)}{\partial\lambda} > 0$ and $\frac{\partial\mathcal{NV}^{pm}(\theta,\lambda)}{\partial\lambda} \le 0$.

Proof of Proposition 6

Let us start by calculating $\mathbb{E}[\theta+q|\theta,\lambda]$. Proposition 2 implies

$$\begin{aligned}
\mathbb{E}[\theta+q|\theta,\lambda] = \theta + \mathbb{E}[q|\theta,\lambda] &= \theta + \mathbb{E}\left[\frac{-\widetilde{r}(\lambda')\theta+\widetilde{r}(\lambda)\theta'}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}|\theta,\lambda\right] \\
&= \theta + \int_0^M\int_{-\infty}^{\infty}\frac{m(\lambda,\lambda')}{m(\lambda,\Lambda)}\frac{-\widetilde{r}(\lambda')\theta+\widetilde{r}(\lambda)\theta'}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}G(d\theta',d\lambda') \\
&= \theta + \int_0^M \frac{m(\lambda,\lambda')}{m(\lambda,\Lambda)}\frac{-\widetilde{r}(\lambda')\theta+\widetilde{r}(\lambda)\mathbb{E}_g[\theta'|\lambda']}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}\psi(\lambda')d\lambda' \\
&= \theta - \frac{\theta}{m(\lambda,\Lambda)}\int_0^M m(\lambda,\lambda')\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}\psi(\lambda')d\lambda' \\
&= \theta - \frac{2(\widetilde{r}(\lambda)-r)}{m(\lambda,\Lambda)}\theta = \theta\left[1-\frac{2(\widetilde{r}(\lambda)-r)}{m(\lambda,\Lambda)}\right],
\end{aligned}$$

where the last equality follows from the definition of $\widetilde{r}(\lambda)$ in Theorem 1 and the previous one follows from the fact that $\mathbb{E}_g[\theta'|\lambda'] = 0$ for $\lambda' \in [0,M]$.

Now, let us calculate $var\left[\theta+q|\theta,\lambda\right]$.

$$
\begin{aligned}
var\left[\theta+q|\theta,\lambda\right] &= \mathbb{E}\left[\left(\theta+q-\mathbb{E}\left[\theta+q|\theta,\lambda\right]\right)^2|\theta,\lambda\right] \\
&= \mathbb{E}\left[\left(\theta+q-\theta\left[1-\frac{2\left(\widetilde{r}\left(\lambda\right)-r\right)}{m\left(\lambda,\Lambda\right)}\right]\right)^2|\theta,\lambda\right] \\
&= \mathbb{E}\left[\left(q+\theta\frac{2\left(\widetilde{r}\left(\lambda\right)-r\right)}{m\left(\lambda,\Lambda\right)}\right)^2|\theta,\lambda\right] \\
&= \mathbb{E}\left[\left(\frac{-\widetilde{r}\left(\lambda'\right)\theta+\widetilde{r}\left(\lambda\right)\theta'}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda'\right)}+\theta\frac{2\left(\widetilde{r}\left(\lambda\right)-r\right)}{m\left(\lambda,\Lambda\right)}\right)^2|\theta,\lambda\right] \\
&= \mathbb{E}\left[\left(\theta\left[\frac{2\left(\widetilde{r}\left(\lambda\right)-r\right)}{m\left(\lambda,\Lambda\right)}-\frac{\widetilde{r}\left(\lambda'\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda'\right)}\right]+\frac{\widetilde{r}\left(\lambda\right)\theta'}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda'\right)}\right)^2|\theta,\lambda\right] \\
&= \mathbb{E}\left[\left(\theta\left[\frac{2\left(\widetilde{r}\left(\lambda\right)-r\right)}{m\left(\lambda,\Lambda\right)}-\frac{\widetilde{r}\left(\lambda'\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda'\right)}\right]\right)^2|\theta,\lambda\right]+\mathbb{E}\left[\left(\frac{\widetilde{r}\left(\lambda\right)\theta'}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda'\right)}\right)^2|\theta,\lambda\right] \\
&= \theta^2 var\left[\frac{\widetilde{r}\left(\lambda'\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda'\right)}\mid\lambda\right] \\
&\qquad +\int_0^M\frac{m\left(\lambda,\lambda'\right)}{m\left(\lambda,\Lambda\right)}var_g\left[\theta\mid\lambda'\right]\left(\frac{\widetilde{r}\left(\lambda\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda'\right)}\right)^2\psi\left(\lambda'\right)d\lambda',
\end{aligned}
$$

where the last equality follows from the definition of $\widetilde{r}\left(\lambda\right)$ in Theorem 1 and the previous one follows from the fact that $\mathbb{E}_g\left[\theta'|\lambda'\right]=0$ for $\lambda'\in[0,M]$.

The definition of $\widetilde{r}\left(\lambda\right)$ in Theorem 1 implies

$$
\frac{2\left(\widetilde{r}\left(\lambda\right)-r\right)}{m\left(\lambda,\Lambda\right)}=\int_0^M\frac{m\left(\lambda,\lambda'\right)}{m\left(\lambda,\Lambda\right)}\frac{\widetilde{r}\left(\lambda'\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda'\right)}\psi(\lambda')d\lambda'\in\left(0,1\right),
$$

because $\widetilde{r}\left(\lambda\right)\geq r$ and

$$
\int_0^M\frac{m\left(\lambda,\lambda'\right)}{m\left(\lambda,\Lambda\right)}\psi(\lambda')d\lambda'=1.
$$

Calculate the derivative of this:

$$
\frac{d}{d\lambda}\frac{2\left(\widetilde{r}\left(\lambda\right)-r\right)}{m\left(\lambda,\Lambda\right)}=\frac{2\widetilde{r}'\left(\lambda\right)m\left(\lambda,\Lambda\right)-2\left(\widetilde{r}\left(\lambda\right)-r\right)m_1\left(\lambda,\Lambda\right)}{\left(m\left(\lambda,\Lambda\right)\right)^2}<0,
$$

which follows by taking the derivative of (3.17) and using the fact that $\widetilde{r}'\left(\lambda\right)>0$.

Lastly, the definition of $\widetilde{r}(\lambda)$ in Theorem 1 implies

$$\begin{aligned} & var\left[\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')} \mid \lambda\right] \\ & \quad = \int_0^M \frac{\lambda'}{\Lambda}\left(\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}\right)^2 \psi(\lambda')d\lambda' - \left(\frac{2\left(\widetilde{r}(\lambda)-r\right)}{m(\lambda,\Lambda)}\right)^2 \in (0,1), \end{aligned}$$

because both terms on the RHS are between 0 and 1 and the first term is larger. Calculate the derivative of this:

$$\begin{aligned} & \frac{d}{d\lambda} var\left[\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')} \mid \lambda\right] \\ & = -2\widetilde{r}'(\lambda)\int_0^M \frac{\lambda'}{\Lambda}\frac{\widetilde{r}(\lambda')}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda')\right)^2}\frac{\widetilde{r}(\lambda')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}\psi(\lambda')d\lambda' \\ & - \frac{8\left(\widetilde{r}(\lambda)-r\right)\widetilde{r}'(\lambda)}{\left(m(\lambda,\Lambda)\right)^2} + \frac{8\left(\left(\widetilde{r}(\lambda)-r\right)\right)^2}{\left(m(\lambda,\Lambda)\right)^2\lambda} \\ & = \frac{4\widetilde{r}'(\lambda)}{m(\lambda,\Lambda)} - \frac{4\left(\widetilde{r}(\lambda)-r\right)}{m(\lambda,\Lambda)\lambda} \\ & - 2\widetilde{r}'(\lambda)\int_0^M \frac{\lambda'}{\Lambda}\frac{\widetilde{r}(\lambda')}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda')\right)^2}\frac{\widetilde{r}(\lambda)}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}\psi(\lambda')d\lambda' \\ & - \frac{8\left(\widetilde{r}(\lambda)-r\right)\widetilde{r}'(\lambda)}{\left(m(\lambda,\Lambda)\right)^2} + \frac{8\left(\left(\widetilde{r}(\lambda)-r\right)\right)^2}{\left(m(\lambda,\Lambda)\right)^2\lambda} \\ & = \frac{4}{m(\lambda,\Lambda)}\left[1-\frac{2\left(\widetilde{r}(\lambda)-r\right)}{m(\lambda,\Lambda)}\right]\left[\widetilde{r}'(\lambda)-\frac{\widetilde{r}(\lambda)-r}{\lambda}\right] \\ & \qquad - 2\widetilde{r}'(\lambda)\int_0^M \frac{\lambda'}{\Lambda}\frac{\widetilde{r}(\lambda')}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda')\right)^2}\frac{\widetilde{r}(\lambda)}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda')}\psi(\lambda')d\lambda' < 0, \end{aligned}$$

where the second equality follows from

$$\frac{2\left(\widetilde{r}(\lambda)-r\right)}{m(\lambda,\Lambda)\lambda} - \frac{2\widetilde{r}'(\lambda)}{m(\lambda,\Lambda)} = \widetilde{r}'(\lambda)\int_0^M \frac{\lambda'}{\Lambda}\frac{\widetilde{r}(\lambda')}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda')\right)^2}\psi(\lambda')d\lambda',$$

which follows by taking the derivative of (3.17).

APPENDIX D: DETAILS OF THE WALRASIAN BENCHMARK

I solve the stationary equilibrium of a continuous frictionless Walrasian market as a benchmark. As is typical in models with continuous access to a trading venue but infrequent need to trade, I start by decomposing the state space into *inaction* and *action* regions. In the inaction region, an investor enjoys the flow utility from holding the asset. In the action region, she immediately accesses the Walrasian market and rebalances her asset position to end up in the inaction region.

The flow Bellman equation of investors in the inaction region can be written as the following integral equation:

$$u(\delta, a) - rJ^W(\delta, a) + \alpha \int_{\delta_L}^{\delta_H} \left[J^W(\delta', a) - J^W(\delta, a)\right] f(\delta')d\delta' = 0. \tag{D.1}$$

The first term is the investor's utility flow. The second term is the time discount. The last term is the expected change in the investor's continuation utility, conditional on switching taste types, which occurs with Poisson intensity α.

In the action region, the value function satisfies the condition

$$J^W(\delta, a) = \max_{\overline{a}} \left\{J^W(\delta, \overline{a}) - P^W(\overline{a} - a)\right\}, \tag{D.2}$$

which basically states that it is indeed optimal for the investor to access the market, costing her $P^W(\overline{a} - a)$ units of the numéraire, where P^W is the market-clearing price. In addition, I need to make sure that staying at a given asset position level in the action region for an infinitesimal amount of time results in a marginal utility loss. Combining with (D.1), this means that $J^W(\delta, a)$ must satisfy the following variational inequality:

$$u(\delta, a) - rJ^W(\delta, a) + \alpha \int_{\delta_L}^{\delta_H} \left[J^W(\delta', a) - J^W(\delta, a)\right] f(\delta')d\delta' \leq 0.$$

Collecting together, the flow Bellman equation of investors can be written as an impulse control problem:

$$\begin{aligned} \max \Big\{ &u(\delta, a) - rJ^W(\delta, a) + \alpha \int_{\delta_L}^{\delta_H} \left[J^W(\delta', a) - J^W(\delta, a)\right] f(\delta')d\delta', \\ &J^W(\delta, a) - \left(J^W(\delta, \overline{a}) - P^W(\overline{a} - a)\right)\Big\} = 0, \end{aligned}$$

where

$$\overline{a} = \operatorname*{argmax}_{\overline{a}} \left\{J^W(\delta, \overline{a}) - P^W(\overline{a} - a)\right\}.$$

Thanks to the absence of frictions, I conjecture (and later verify) that, given P^W, the inaction region is a measure-zero point $\left[\delta, \hat{a}(\delta; P^W)\right]$ for investors with taste type δ, where $\hat{a}(.;P^W)$ is a strictly monotone function. Under this conjecture, one can use (D.2) to substitute out $J^W(\delta', a)$ in (D.1) to obtain the auxiliary HJB equation (3.8) of Subsection 3.2.

$$rJ^W(\delta, a) = u(\delta, a) + \alpha \int_{\delta_L}^{\delta_H} \max_{a'} \left\{ J^W(\delta', a') - J^W(\delta, a) - P^W(a' - a) \right\} f(\delta') d\delta'.$$

The FOC for the asset position and the envelope condition[3] are

$$J_2^W(\delta', a') = P^W$$

and

$$rJ_2^W(\delta, a) = u_2(\delta, a) + \alpha\left(-J_2^W(\delta, a) + P^W\right),$$

where $u_2(.,.)$ represents the partial derivative with respect to the second argument. Combining these two conditions, I get the optimal demand of the investor with δ, which places her in the inaction region:

$$\hat{a}(\delta; P^W) = \frac{r}{\kappa}\left(\frac{\delta}{r} - P^W\right).$$

APPENDIX E: CALCULATION OF INTERMEDIATION MARKUPS

First, calculate the transaction price for the initial trade at which the investor with 0 inventory and speed type λ provides intermediation to a counterparty with speed type λ' by buying θ units of the asset from him. According to Equation (3.19) this price must be

$$P = J_\theta(\theta, \lambda) + \frac{\kappa\theta}{4}\left(\frac{1}{\widetilde{r}(\lambda)} - \frac{1}{\widetilde{r}(\lambda')}\right).$$

Using the marginal valuation formula from Proposition 2,

(E.1)

$$P = \underbrace{\frac{u_2(\overline{\delta}, A)}{r} - \kappa\frac{\theta}{\widetilde{r}(\lambda)}}_{P^{ihr}} + \underbrace{\frac{\kappa}{4}\left(\frac{1}{\widetilde{r}(\lambda)} - \frac{1}{\widetilde{r}(\lambda')}\right)\theta}_{P^{sp}} = \frac{u_2(\overline{\delta}, A)}{r} - \frac{\kappa\theta}{4}\left(\frac{3}{\widetilde{r}(\lambda)} + \frac{1}{\widetilde{r}(\lambda')}\right),$$

where P^{ihr} is the post-trade marginal valuation and P^{sp} is the speed premium.

Now, calculate the expected price the investor will receive while trying to unload this inventory of θ:

$$\frac{\mathbb{E}[Pq|\theta, \lambda; \eta]}{\mathbb{E}[q|\theta, \lambda; \eta]}.$$

[3]To write down these conditions, I assume that $J^W(\delta, .)$ is strictly concave and continuously differentiable. This assumption is also verified *ex post*.

Let us start by calculating $\mathbb{E}\left[q|\theta,\lambda;\eta\right]$. First, note that since δ is distributed symmetrically, Equation (3.28) implies $\widehat{g}_{\lambda''}(z)=\widehat{g}_{\lambda''}(-z)$, and hence, θ is distributed symmetrically conditional on λ''. Then, $\mathbb{E}_g\left[\theta''|\lambda'';\eta\right]=0$ for $\lambda''\in[0,M]$. Proposition 2 implies

$$\begin{aligned}
\mathbb{E}\left[q|\theta,\lambda;\eta\right] &= \mathbb{E}\left[\frac{-\widetilde{r}\left(\lambda''\right)\theta+\widetilde{r}\left(\lambda\right)\theta''}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}|\theta,\lambda;\eta\right] \\
&= \int_0^M\int_{-\eta}^{\eta}\frac{m\left(\lambda,\lambda''\right)}{m\left(\lambda,\Lambda\right)}\frac{-\widetilde{r}\left(\lambda''\right)\theta+\widetilde{r}\left(\lambda\right)\theta''}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\frac{G\left(d\theta'',d\lambda''\right)}{G_{\lambda''}\left(\eta\right)-G_{\lambda''}\left(-\eta\right)} \\
&= \int_0^M\frac{m\left(\lambda,\lambda''\right)}{m\left(\lambda,\Lambda\right)}\frac{-\widetilde{r}\left(\lambda''\right)\theta+\widetilde{r}\left(\lambda\right)\mathbb{E}_g\left[\theta''|\lambda'';\eta\right]}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\psi\left(\lambda''\right)d\lambda'' \\
&= -\frac{\theta}{m\left(\lambda,\Lambda\right)}\int_0^M m\left(\lambda,\lambda''\right)\frac{\widetilde{r}\left(\lambda''\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\psi\left(\lambda''\right)d\lambda'' \\
&= -\frac{2\left(\widetilde{r}\left(\lambda\right)-r\right)}{m\left(\lambda,\Lambda\right)}\theta,
\end{aligned} \tag{E.2}$$

where the last equality follows from the definition of $\widetilde{r}\left(\lambda\right)$ in Theorem 1 and the previous one follows from the fact that $\mathbb{E}_g\left[\theta''|\lambda'';\eta\right]=0$ for $\lambda''\in[0,M]$.

Now, let us calculate $\mathbb{E}\left[Pq|\theta,\lambda;\eta\right]$. $\mathbb{E}\left[Pq|\theta,\lambda;\eta\right]$ will have a component due to post-trade marginal valuation and another component due to speed premium. Call these, respectively, $\mathbb{E}^{ihr}\left[Pq|\theta,\lambda;\eta\right]$ and $\mathbb{E}^{sp}\left[Pq|\theta,\lambda;\eta\right]$. First, note from Proposition 2 that the transaction price $P\left[\left(\theta,\lambda\right),\left(\theta'',\lambda''\right)\right]$ can be written as

$$\underbrace{\frac{u_2\left(\overline{\delta},A\right)}{r}-\kappa\frac{\theta+\theta''}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}}_{\text{post-trade marg. val.}}+\underbrace{\frac{\kappa}{4}\frac{\widetilde{r}\left(\lambda''\right)-\widetilde{r}\left(\lambda\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\left(-\frac{\theta}{\widetilde{r}\left(\lambda\right)}+\frac{\theta''}{\widetilde{r}\left(\lambda''\right)}\right)}_{\text{speed premium}}.$$

Thus,

$$\begin{aligned}
&\mathbb{E}^{ihr}\left[Pq|\theta,\lambda;\eta\right] \\
&= \mathbb{E}\left[\frac{-\widetilde{r}\left(\lambda''\right)\theta+\widetilde{r}\left(\lambda\right)\theta''}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\left(\frac{u_2\left(\overline{\delta},A\right)}{r}-\kappa\frac{\theta+\theta''}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\right)|\theta,\lambda;\eta\right]
\end{aligned}$$

$$= \mathbb{E}\left[\frac{-\widetilde{r}(\lambda'')\theta + \widetilde{r}(\lambda)\theta''}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')}\frac{u_2\left(\overline{\delta}, A\right)}{r} + \kappa\theta^2\frac{\widetilde{r}(\lambda'')}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')\right)^2}\right.$$
$$\left. -\kappa\left(\theta''\right)^2\frac{\widetilde{r}(\lambda)}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')\right)^2}|\theta,\lambda;\eta\right]$$
$$= \mathbb{E}\left[q|\theta,\lambda;\eta\right]\frac{u_2\left(\overline{\delta}, A\right)}{r} + \kappa\theta^2\int_0^M \frac{m(\lambda,\lambda'')}{m(\lambda,\Lambda)}\frac{\widetilde{r}(\lambda'')}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')\right)^2}\psi(\lambda'')\,d\lambda''$$
$$-\kappa\int_0^M \frac{m(\lambda,\lambda'')}{m(\lambda,\Lambda)}\frac{\widetilde{r}(\lambda)}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')\right)^2}\mathbb{E}_g\left[\left(\theta''\right)^2|\lambda'';\eta\right]\psi(\lambda'')\,d\lambda'',$$

where the last equality follows from (E.2) and the previous equality follows from the fact that $\mathbb{E}_g[\theta''|\lambda'';\eta] = 0$ for $\lambda'' \in [0, M]$. Similarly,

$$\mathbb{E}^{sp}\left[Pq|\theta,\lambda;\eta\right] =$$
$$\mathbb{E}\left[\frac{-\widetilde{r}(\lambda'')\theta + \widetilde{r}(\lambda)\theta''}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')}\left\{\frac{\kappa}{4}\frac{\widetilde{r}(\lambda'')-\widetilde{r}(\lambda)}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')}\left(-\frac{\theta}{\widetilde{r}(\lambda)}+\frac{\theta''}{\widetilde{r}(\lambda'')}\right)\right\}|\theta,\lambda;\eta\right]$$
$$= \mathbb{E}\left[\frac{\kappa\theta^2}{4}\frac{\widetilde{r}(\lambda'')-\widetilde{r}(\lambda)}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')\right)^2}\frac{\widetilde{r}(\lambda'')}{\widetilde{r}(\lambda)} + \frac{\kappa\left(\theta''\right)^2}{4}\frac{\widetilde{r}(\lambda'')-\widetilde{r}(\lambda)}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')\right)^2}\frac{\widetilde{r}(\lambda)}{\widetilde{r}(\lambda'')}|\theta,\lambda;\eta\right]$$
$$= \frac{\kappa\theta^2}{4}\int_0^M \frac{m(\lambda,\lambda'')}{m(\lambda,\Lambda)}\frac{\widetilde{r}(\lambda'')-\widetilde{r}(\lambda)}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')\right)^2}\frac{\widetilde{r}(\lambda'')}{\widetilde{r}(\lambda)}\psi(\lambda'')\,d\lambda''$$
$$+\frac{\kappa}{4}\int_0^M \frac{m(\lambda,\lambda'')}{m(\lambda,\Lambda)}\frac{\widetilde{r}(\lambda'')-\widetilde{r}(\lambda)}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')\right)^2}\frac{\widetilde{r}(\lambda)}{\widetilde{r}(\lambda'')}\mathbb{E}_g\left[\left(\theta''\right)^2|\lambda'';\eta\right]\psi(\lambda'')\,d\lambda''.$$

Then, the expected price the investor will receive by unloading the inventory of θ becomes:

$$\frac{\mathbb{E}\left[Pq|\theta,\lambda;\eta\right]}{\mathbb{E}\left[q|\theta,\lambda;\eta\right]} = \frac{\mathbb{E}^{ihr}\left[Pq|\theta,\lambda;\eta\right]}{\mathbb{E}\left[q|\theta,\lambda;\eta\right]} + \frac{\mathbb{E}^{sp}\left[Pq|\theta,\lambda;\eta\right]}{\mathbb{E}\left[q|\theta,\lambda;\eta\right]}, \tag{E.3}$$

where

$$\frac{\mathbb{E}^{ihr}\left[Pq|\theta,\lambda;\eta\right]}{\mathbb{E}\left[q|\theta,\lambda;\eta\right]} = \frac{u_2\left(\overline{\delta}, A\right)}{r}$$
$$-\kappa\theta\int_0^M \frac{m(\lambda,\lambda'')}{2\left(\widetilde{r}(\lambda)-r\right)}\frac{\widetilde{r}(\lambda'')}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')\right)^2}\psi(\lambda'')\,d\lambda''$$
$$+\frac{\kappa}{\theta}\int_0^M \frac{m(\lambda,\lambda'')}{2\left(\widetilde{r}(\lambda)-r\right)}\frac{\widetilde{r}(\lambda)}{\left(\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')\right)^2}var_g\left[\theta''|\lambda'';\eta\right]\psi(\lambda'')\,d\lambda''$$

and

$$\frac{\mathbb{E}^{sp}\left[Pq|\theta,\lambda;\eta\right]}{\mathbb{E}\left[q|\theta,\lambda;\eta\right]}=\frac{\kappa\theta}{4}\int_0^M\frac{m\left(\lambda,\lambda''\right)}{2\left(\widetilde{r}\left(\lambda\right)-r\right)}\frac{\widetilde{r}\left(\lambda\right)-\widetilde{r}\left(\lambda''\right)}{\left(\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)\right)^2}\frac{\widetilde{r}\left(\lambda''\right)}{\widetilde{r}\left(\lambda\right)}\psi\left(\lambda''\right)d\lambda''$$

$$+\frac{\kappa}{4\theta}\int_0^M\frac{m\left(\lambda,\lambda''\right)}{2\left(\widetilde{r}\left(\lambda\right)-r\right)}\frac{\widetilde{r}\left(\lambda\right)-\widetilde{r}\left(\lambda''\right)}{\left(\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)\right)^2}\frac{\widetilde{r}\left(\lambda\right)}{\widetilde{r}\left(\lambda''\right)}var_g\left[\theta''|\lambda'';\eta\right]\psi\left(\lambda''\right)d\lambda''.$$

Define the markup as

$$\mu\left(\theta,\lambda,\lambda'\right)\equiv\frac{\frac{\mathbb{E}[Pq|\theta,\lambda;\eta]}{\mathbb{E}[q|\theta,\lambda;\eta]}-P}{P}=\underbrace{\frac{\frac{\mathbb{E}^{ihr}[Pq|\theta,\lambda;\eta]}{\mathbb{E}[q|\theta,\lambda;\eta]}-P^{ihr}}{P}}_{\equiv\mu^{ihr}(\theta,\lambda,\lambda')}+\underbrace{\frac{\frac{\mathbb{E}^{sp}[Pq|\theta,\lambda;\eta]}{\mathbb{E}[q|\theta,\lambda;\eta]}-P^{sp}}{P}}_{\equiv\mu^{sp}(\theta,\lambda,\lambda')}.$$

Using (E.1), (E.3), and the fact that

$$2\left(\widetilde{r}\left(\lambda\right)-r\right)=\int_0^M m\left(\lambda,\lambda''\right)\frac{\widetilde{r}\left(\lambda''\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\psi\left(\lambda''\right)d\lambda'',$$

one obtains (4.6).

Using the same equation and the fact that $\widetilde{r}\left(\lambda\right)\geq r$ for all $\lambda\in[0,M]$, one can also show that the markup (4.6) is positive when the normalizing price (E.1) and θ are positive.

Proof of Proposition 7

Rewrite the numerator of markup:

$$\frac{\kappa}{4\theta}\int_0^M\frac{m\left(\lambda,\lambda''\right)}{2\left(\widetilde{r}\left(\lambda\right)-r\right)}\frac{\widetilde{r}\left(\lambda\right)}{\left(\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)\right)^2}\left[\frac{\widetilde{r}\left(\lambda\right)}{\widetilde{r}\left(\lambda''\right)}+3\right]var_g\left[\theta''|\lambda'';\eta\right]\psi\left(\lambda''\right)d\lambda''+\epsilon\left(\lambda\right),$$

where $\epsilon\left(\lambda\right)$ collects the terms that do not contain $var_g\left[\theta''|\lambda'';\eta\right]$. Take derivative w.r.t. λ:

$$\frac{\kappa}{4\theta}\int_0^M \frac{m_\lambda(\lambda,\lambda'')\,2\,(\widetilde{r}(\lambda)-r)-m(\lambda,\lambda'')\,2\widetilde{r}'(\lambda)}{\left[2\,(\widetilde{r}(\lambda)-r)\right]^2}$$

$$\frac{[\widetilde{r}(\lambda)]^2+3\widetilde{r}(\lambda'')\,\widetilde{r}(\lambda)}{(\widetilde{r}(\lambda)+\widetilde{r}(\lambda''))^2}\frac{1}{\widetilde{r}(\lambda'')}var\left[\theta''|\lambda'';\eta\right]\psi(\lambda'')\,d\lambda''$$

$$+\frac{\kappa}{4\theta}\int_0^M \frac{m(\lambda,\lambda'')}{2\,(\widetilde{r}(\lambda)-r)}\widetilde{r}'(\lambda)\,\widetilde{r}(\lambda'')\,\frac{3\widetilde{r}(\lambda'')-\widetilde{r}(\lambda)}{(\widetilde{r}(\lambda)+\widetilde{r}(\lambda''))^3}$$

$$\frac{1}{\widetilde{r}(\lambda'')}var\left[\theta''|\lambda'';\eta\right]\psi(\lambda'')\,d\lambda''+\epsilon'(\lambda)$$

$$=\frac{\kappa}{4\theta}\int_0^M \frac{m(\lambda,\lambda'')}{2\,(\widetilde{r}(\lambda)-r)}\frac{[\widetilde{r}(\lambda)]^2+3\widetilde{r}(\lambda'')\,\widetilde{r}(\lambda)}{(\widetilde{r}(\lambda)+\widetilde{r}(\lambda''))^2}\frac{1}{\lambda}\frac{1}{\widetilde{r}(\lambda'')}var\left[\theta''|\lambda'';\eta\right]\psi(\lambda'')d\lambda''$$

$$-\frac{\kappa}{4\theta}\int_0^M \frac{m(\lambda,\lambda'')}{2\,(\widetilde{r}(\lambda)-r)}\frac{[\widetilde{r}(\lambda)]^2+3\widetilde{r}(\lambda'')\,\widetilde{r}(\lambda)}{(\widetilde{r}(\lambda)+\widetilde{r}(\lambda''))^2}\frac{\widetilde{r}'(\lambda)}{\widetilde{r}(\lambda)-r}$$

$$\frac{1}{\widetilde{r}(\lambda'')}var\left[\theta''|\lambda'';\eta\right]\psi(\lambda'')\,d\lambda''$$

$$-\frac{\kappa}{4\theta}\int_0^M \frac{m(\lambda,\lambda'')}{2\,(\widetilde{r}(\lambda)-r)}\frac{\widetilde{r}(\lambda)\,\widetilde{r}(\lambda'')}{(\widetilde{r}(\lambda)+\widetilde{r}(\lambda''))^2}\frac{\widetilde{r}'(\lambda)}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')}$$

$$\frac{1}{\widetilde{r}(\lambda'')}var\left[\theta''|\lambda'';\eta\right]\psi(\lambda'')\,d\lambda''$$

$$+\frac{\kappa}{4\theta}\int_0^M \frac{m(\lambda,\lambda'')}{2\,(\widetilde{r}(\lambda)-r)}\widetilde{r}'(\lambda)\,\widetilde{r}(\lambda'')\,\frac{3\widetilde{r}(\lambda'')}{(\widetilde{r}(\lambda)+\widetilde{r}(\lambda''))^3}$$

$$\frac{1}{\widetilde{r}(\lambda'')}var\left[\theta''|\lambda'';\eta\right]\psi(\lambda'')\,d\lambda''+\epsilon'(\lambda)\,.$$

Using the fact that

$$\widetilde{r}'(\lambda)\left[1+\int_0^M m(\lambda,\lambda'')\,\frac{\widetilde{r}(\lambda'')}{(\widetilde{r}(\lambda)+\widetilde{r}(\lambda''))^2}\psi(\lambda'')\,d\lambda''\right]=\frac{\widetilde{r}(\lambda)-r}{\lambda}, \tag{E.4}$$

one can show that the sum of the terms before $\epsilon'(\lambda)$ is positive. It is easy to verify that the denominator of markup is an increasing function of λ, and hence, it will contribute negatively to the derivative of markup. Also, the sign of $\epsilon'(\lambda)$ can be negative or positive. However, it is certain that the terms with

$var_g\left[\theta''|\lambda'';\eta\right]$ contribute positively to the derivative of markup. Thus, from continuity, $var_g\left[\theta''|\lambda'';\eta\right]$s must be large enough for the total derivative to be positive, which completes the part (ii) of the proposition.

To show the part (i), rewrite the numerator of markup:

$$\frac{\kappa\theta}{4\widetilde{r}\left(\lambda'\right)}+\frac{1}{2}\frac{\kappa\theta}{\widetilde{r}\left(\lambda\right)}\int_0^M\frac{m\left(\lambda,\lambda''\right)}{2\left(\widetilde{r}\left(\lambda\right)-r\right)}\left(\frac{\widetilde{r}\left(\lambda''\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\right)^2\psi\left(\lambda''\right)d\lambda''+\varepsilon\left(\lambda\right),$$

where $\varepsilon\left(\lambda\right)$ represents the terms with integral of $var_g\left[\theta''|\lambda'';\eta\right]$.

Take the derivative w.r.t. λ:

$$\begin{aligned}
&-\frac{1}{2}\frac{\kappa\theta\widetilde{r}'\left(\lambda\right)}{\left[\widetilde{r}\left(\lambda\right)\right]^2}\int_0^M\frac{m\left(\lambda,\lambda''\right)}{2\left(\widetilde{r}\left(\lambda\right)-r\right)}\left(\frac{\widetilde{r}\left(\lambda''\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\right)^2\psi\left(\lambda''\right)d\lambda''\\
&+\frac{1}{2}\frac{\kappa\theta}{\widetilde{r}\left(\lambda\right)}\int_0^M\frac{m_\lambda\left(\lambda,\lambda''\right)2\left(\widetilde{r}\left(\lambda\right)-r\right)-m\left(\lambda,\lambda''\right)2\widetilde{r}'\left(\lambda\right)}{\left[2\left(\widetilde{r}\left(\lambda\right)-r\right)\right]^2}\\
&\left(\frac{\widetilde{r}\left(\lambda''\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\right)^2\psi\left(\lambda''\right)d\lambda''-\frac{1}{2}\frac{\kappa\theta}{\widetilde{r}\left(\lambda\right)}\int_0^M\frac{m\left(\lambda,\lambda''\right)}{2\left(\widetilde{r}\left(\lambda\right)-r\right)}2\left(\frac{\widetilde{r}\left(\lambda''\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\right)\\
&\frac{\widetilde{r}\left(\lambda''\right)\widetilde{r}'\left(\lambda\right)}{\left(\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)\right)^2}\psi\left(\lambda''\right)d\lambda''+\varepsilon'\left(\lambda\right)\\
&=-\frac{1}{2}\frac{\kappa\theta\widetilde{r}'\left(\lambda\right)}{\widetilde{r}\left(\lambda\right)}\int_0^M\frac{m\left(\lambda,\lambda''\right)}{2\left(\widetilde{r}\left(\lambda\right)-r\right)}\left(\frac{\widetilde{r}\left(\lambda''\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\right)^2\frac{1}{\widetilde{r}\left(\lambda\right)}\psi\left(\lambda''\right)d\lambda''\\
&+\frac{1}{2}\frac{\kappa\theta}{\widetilde{r}\left(\lambda\right)}\int_0^M\frac{m\left(\lambda,\lambda''\right)2\left(\widetilde{r}\left(\lambda\right)-r\right)-m\left(\lambda,\lambda''\right)\lambda2\widetilde{r}'\left(\lambda\right)}{\left[2\left(\widetilde{r}\left(\lambda\right)-r\right)\right]^2\lambda}\\
&\left(\frac{\widetilde{r}\left(\lambda''\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\right)^2\psi\left(\lambda''\right)d\lambda''-\frac{1}{2}\frac{\kappa\theta\widetilde{r}'\left(\lambda\right)}{\widetilde{r}\left(\lambda\right)}\int_0^M\frac{m\left(\lambda,\lambda''\right)}{2\left(\widetilde{r}\left(\lambda\right)-r\right)}\\
&\left(\frac{\widetilde{r}\left(\lambda''\right)}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\right)^2\frac{2}{\widetilde{r}\left(\lambda\right)+\widetilde{r}\left(\lambda''\right)}\psi\left(\lambda''\right)d\lambda''+\varepsilon'\left(\lambda\right)
\end{aligned}$$

$$= \frac{1}{2}\frac{\kappa\theta\widetilde{r}'(\lambda)}{\widetilde{r}(\lambda)}\int_0^M \frac{m(\lambda,\lambda'')}{2(\widetilde{r}(\lambda)-r)}\left(\frac{\widetilde{r}(\lambda'')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')}\right)^2 \frac{1}{\lambda\widetilde{r}'(\lambda)}\psi(\lambda'')\,d\lambda''$$

$$-\frac{1}{2}\frac{\kappa\theta\widetilde{r}'(\lambda)}{\widetilde{r}(\lambda)}\int_0^M \frac{m(\lambda,\lambda'')}{2(\widetilde{r}(\lambda)-r)}\left(\frac{\widetilde{r}(\lambda'')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')}\right)^2 \frac{1}{\widetilde{r}(\lambda)}\psi(\lambda'')\,d\lambda''$$

$$-\frac{1}{2}\frac{\kappa\theta\widetilde{r}'(\lambda)}{\widetilde{r}(\lambda)}\int_0^M \frac{m(\lambda,\lambda'')}{2(\widetilde{r}(\lambda)-r)}\left(\frac{\widetilde{r}(\lambda'')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')}\right)^2 \frac{2}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')}\psi(\lambda'')\,d\lambda''$$

$$-\frac{1}{2}\frac{\kappa\theta\widetilde{r}'(\lambda)}{\widetilde{r}(\lambda)}\int_0^M \frac{m(\lambda,\lambda'')}{2(\widetilde{r}(\lambda)-r)}\left(\frac{\widetilde{r}(\lambda'')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')}\right)^2 \frac{1}{\widetilde{r}(\lambda)-r}\psi(\lambda'')\,d\lambda''+\varepsilon'(\lambda)$$

$$= \frac{1}{2}\frac{\kappa\theta\widetilde{r}'(\lambda)}{\widetilde{r}(\lambda)}\int_0^M \frac{m(\lambda,\lambda'')}{2(\widetilde{r}(\lambda)-r)}\left(\frac{\widetilde{r}(\lambda'')}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')}\right)^2$$

$$\left[\frac{1}{\lambda\widetilde{r}'(\lambda)}-\frac{1}{\widetilde{r}(\lambda)}-\frac{2}{\widetilde{r}(\lambda)+\widetilde{r}(\lambda'')}-\frac{1}{\widetilde{r}(\lambda)-r}\right]\psi(\lambda'')\,d\lambda''+\varepsilon'(\lambda).$$

Again, using (E.4) and that the lower bound of the distribution of λs is $1/8$, one can show that the first term of the derivative is negative. Since $\varepsilon'(\lambda)$ is positive, from continuity, $var_g[\theta''|\lambda'';\eta]$s must be small enough for the total derivative to be negative. It is easy to verify that the denominator of markup is an increasing function of λ. Thus, the derivative of the markup is negative when $var_g[\theta''|\lambda'';\eta]$s are small enough.

APPENDIX F: PLANNER'S PROBLEM

I define social welfare as the discounted sum of the utility flows of all investors,

$$\mathbb{W} = \int_0^\infty e^{-rt}\left\{\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} u(\delta,a)\,\phi_t(\delta,a,\lambda)\,d\delta da d\lambda\right\}dt. \tag{F.1}$$

Any transfer of the numéraire good from one investor to another does not enter $\mathbb{W}$ because of quasi-linear preferences. The planner maximizes $\mathbb{W}$ with respect to controls, $q_t[(\delta,a,\lambda),(\delta',a',\lambda')]$, subject to the laws of motion for the state variables, $\phi_t(\delta,a,\lambda)$, and to the feasibility condition of asset reallocation,

$$q_t[(\delta,a,\lambda),(\delta',a',\lambda')]+q_t[(\delta',a',\lambda'),(\delta,a,\lambda)]=0, \tag{F.2}$$

which also results in the imposition that the solution does not depend on the identities or "names" of investors.

Since δ, a, and λ are continuous variables, we have a continuum of control variables (and of dynamic restrictions and co-state variables, too), corresponding to the continuum of investor characteristics. van Imhoff (1982) describes a heuristic method of solving such problems. This method relies on interpreting the integral (F.1) as a summation of discrete variables over intervals with widths $d\delta$, da, and $d\lambda$. An application of *Lebesgue dominated convergence theorem*[4] guarantees the convergence of this summation to the integral (F.1) as the widths of intervals approach 0.

Keeping in mind van Imhoff (1982)'s interpretation, the planner's current-value Hamiltonian can be written as

$$
\begin{aligned}
L(q|\Phi) = & \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} u(\delta, a)\, \Phi(d\delta, da, d\lambda) \\
& + \alpha \int_{\delta_L}^{\delta_H} \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \left(\vartheta(\delta', a, \lambda) - \vartheta(\delta, a, \lambda)\right) f(\delta')\, d\delta' \Phi(d\delta, da, d\lambda) \\
& + \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} m(\lambda, \lambda') \{\vartheta(\delta, a + q[(\delta, a, \lambda), (\delta', a', \lambda')], \lambda) \\
& - \vartheta(\delta, a, \lambda)\}\, \Phi(d\delta', da', d\lambda')\, \Phi(d\delta, da, d\lambda) \\
& + \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \zeta[(\delta, a, \lambda), (\delta', a', \lambda')]\, \{q[(\delta, a, \lambda), (\delta', a', \lambda')] \\
& \qquad + q[(\delta', a', \lambda'), (\delta, a, \lambda)]\}\, \Phi(d\delta', da', d\lambda')\, \Phi(d\delta, da, d\lambda),
\end{aligned}
$$

where ϕ induces the cdf Φ; ϑ denotes the current-value co-state variable associated with ϕ; and ζ is the Lagrange multiplier associated with the condition (F.2).

First-order conditions. Take any optimal q^e and let

$$
\vartheta^e(\delta, a, \lambda) = \vartheta(\delta, a + q^e[(\delta, a, \lambda), (\delta', a', \lambda')], \lambda), \tag{F.3}
$$

and let

$$
\begin{aligned}
\hat{q}[(\delta, a, \lambda), (\delta', a', \lambda')] = & \; q^e[(\delta, a, \lambda), (\delta', a', \lambda')] \\
& + \varepsilon \mathbb{I}_{\{\vartheta^e(\delta, a, \lambda) > \vartheta^e(\delta', a', \lambda')\}} - \varepsilon \mathbb{I}_{\{\vartheta^e(\delta, a, \lambda) < \vartheta^e(\delta', a', \lambda')\}} \\
= & \; q^e[(\delta, a, \lambda), (\delta', a', \lambda')] + \varepsilon \Delta[(\delta, a, \lambda), (\delta', a', \lambda')].
\end{aligned}
$$

[4]See, for a reference, Hutson, Pym, and Cloud (2005, p. 55).

For small ε, I obtain up to second-order terms:

$$L\left(\hat{q}|\Phi\right)-L\left(q^{e}|\Phi\right)=\varepsilon\int_{0}^{M}\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_{0}^{M}\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}m\left(\lambda,\lambda'\right)\vartheta_2^{e}\left(\delta,a,\lambda\right)$$
$$\Delta\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]\Phi\left(d\delta',da',d\lambda'\right)\Phi\left(d\delta,da,d\lambda\right)$$
$$+\,\varepsilon\int_{0}^{M}\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_{0}^{M}\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\zeta\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]\left\{\Delta\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]\right.$$
$$\left.+\Delta\left[\left(\delta',a',\lambda'\right),\left(\delta,a,\lambda\right)\right]\right\}\Phi\left(d\delta',da',d\lambda'\right)\Phi\left(d\delta,da,d\lambda\right)$$
$$=\frac{\varepsilon}{2}\int_{0}^{M}\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_{0}^{M}\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}m\left(\lambda,\lambda'\right)\vartheta_2^{e}\left(\delta,a,\lambda\right)\Delta\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]$$
$$\Phi\left(d\delta',da',d\lambda'\right)\Phi\left(d\delta,da,d\lambda\right)+\frac{\varepsilon}{2}\int_{0}^{M}\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_{0}^{M}\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}m\left(\lambda,\lambda'\right)\vartheta_2^{e}\left(\delta',a',\lambda'\right)$$
$$\Delta\left[\left(\delta',a',\lambda'\right),\left(\delta,a,\lambda\right)\right]\Phi\left(d\delta',da',d\lambda'\right)\Phi\left(d\delta,da,d\lambda\right)$$
$$+\,\varepsilon\int_{0}^{M}\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_{0}^{M}\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\zeta\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]\left\{\Delta\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]\right.$$
$$\left.+\Delta\left[\left(\delta',a',\lambda'\right),\left(\delta,a,\lambda\right)\right]\right\}\Phi\left(d\delta',da',d\lambda'\right)\Phi\left(d\delta,da,d\lambda\right)$$
$$=\frac{\varepsilon}{2}\int_{0}^{M}\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_{0}^{M}\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}m\left(\lambda,\lambda'\right)\left\{\vartheta_2^{e}\left(\delta,a,\lambda\right)-\vartheta_2^{e}\left(\delta',a',\lambda'\right)\right\}$$
$$\Delta\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]\Phi\left(d\delta',da',d\lambda'\right)\Phi\left(d\delta,da,d\lambda\right)$$
$$+\,\varepsilon\int_{0}^{M}\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_{0}^{M}\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\zeta\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]\left\{\Delta\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]\right.$$
$$\left.+\Delta\left[\left(\delta',a',\lambda'\right),\left(\delta,a,\lambda\right)\right]\right\}\Phi\left(d\delta',da',d\lambda'\right)\Phi\left(d\delta,da,d\lambda\right)$$

If q^e is optimal, this must be negative. The second term is 0 by construction. Since the integrand in the first term is positive, it must be zero everywhere. Recalling (F.2) and (F.3), thus, the FOC becomes

(F.4)
$$\vartheta_2\left(\delta,a+q^{e}\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right],\lambda\right)=\vartheta_2\left(\delta',a'-q^{e}\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right],\lambda'\right).$$

ODE for co-state variables. In an optimum, the co-state variables must satisfy the ODEs,

$$\nabla_{n(\delta,a,\lambda)} L\left(q^e|\Phi\right) = r\vartheta\left(\delta,a,\lambda\right) - \dot{\vartheta}\left(\delta,a,\lambda\right), \tag{F.5}$$

where $n\left(\delta,a,\lambda\right)$ is the degenerate measure which puts all the probability on the type $\left(\delta,a,\lambda\right)$ and ∇_n denotes the Gâteaux differential in the direction of measure n:

$$\nabla_n L\left(q^e|\Phi\right) = \lim_{\varepsilon\to 0} \frac{L\left(q^e|\Phi+\varepsilon n\right) - L\left(q^e|\Phi\right)}{\varepsilon}.$$

For small ε, I obtain up to second-order terms:

$$\begin{aligned}
&L\left(q^e|\Phi+\varepsilon n\right) - L\left(q^e|\Phi\right) = \varepsilon\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} u\left(\delta,a\right) n\left(d\delta,da,d\lambda\right)\\
&+\varepsilon\alpha\int_{\delta_L}^{\delta_H}\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\left(\vartheta\left(\delta',a,\lambda\right)-\vartheta\left(\delta,a,\lambda\right)\right) f\left(\delta'\right) d\delta' n\left(d\delta,da,d\lambda\right)\\
&+\varepsilon\int_0^M\int_{-\infty}^{\infty}\int_{-1}^{1}\int_0^M\int_{-\infty}^{\infty}\int_{-1}^{1} m\left(\lambda,\lambda'\right)\left\{\vartheta\left(\delta,a+q^e\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right],\lambda\right)\right.\\
&\left.-\vartheta\left(\delta,a,\lambda\right)\right\}\Phi\left(d\delta',da',d\lambda'\right) n\left(d\delta,da,d\lambda\right)\\
&+\varepsilon\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m\left(\lambda,\lambda'\right)\left\{\vartheta\left(\delta,a+q^e\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right],\lambda\right)\right.\\
&\left.-\vartheta\left(\delta,a,\lambda\right)\right\} n\left(d\delta',da',d\lambda'\right)\Phi\left(d\delta,da,d\lambda\right)\\
&+\varepsilon\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\zeta\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]\left\{q^e\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]\right.\\
&\left.+q^e\left[\left(\delta',a',\lambda'\right),\left(\delta,a,\lambda\right)\right]\right\}\Phi\left(d\delta',da',d\lambda'\right) n\left(d\delta,da,d\lambda\right)\\
&+\varepsilon\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\zeta\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]\left\{q^e\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]\right.\\
&\qquad\qquad\left.+q^e\left[\left(\delta',a',\lambda'\right),\left(\delta,a,\lambda\right)\right]\right\} n\left(d\delta',da',d\lambda'\right)\Phi\left(d\delta,da,d\lambda\right).
\end{aligned}$$

Thus,

$$
\begin{aligned}
&\nabla_{n(\delta,a,\lambda)}L\left(q^e|\Phi\right)=u\left(\delta,a\right)+\alpha\int_{\delta_L}^{\delta_H}\left(\vartheta\left(\delta',a,\lambda\right)-\vartheta\left(\delta,a,\lambda\right)\right)f\left(\delta'\right)d\delta'\\
&+\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}m\left(\lambda,\lambda'\right)\left\{\vartheta\left(\delta,a+q^e\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right],\lambda\right)-\vartheta\left(\delta,a,\lambda\right)\right.\\
&\left.+\vartheta\left(\delta',a'+q^e\left[\left(\delta',a',\lambda'\right),\left(\delta,a,\lambda\right)\right],\lambda'\right)-\vartheta\left(\delta',a',\lambda'\right)\right\}\Phi\left(d\delta',da',d\lambda'\right)\\
&+\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\left\{\zeta\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]+\zeta\left[\left(\delta',a',\lambda'\right),\left(\delta,a,\lambda\right)\right]\right\}\\
&\quad\left\{q^e\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right]+q^e\left[\left(\delta',a',\lambda'\right),\left(\delta,a,\lambda\right)\right]\right\}\Phi\left(d\delta',da',d\lambda'\right).
\end{aligned}
$$

Using (F.2), (F.5), and the FOC (F.4), the following ODE for the co-state variables obtains in an optimum:

$$
\begin{aligned}
&r\vartheta\left(\delta,a,\lambda\right)-\dot{\vartheta}\left(\delta,a,\lambda\right)=u\left(\delta,a\right)\\
&+\alpha\int_{\delta_L}^{\delta_H}\left(\vartheta\left(\delta',a,\lambda\right)-\vartheta\left(\delta,a,\lambda\right)\right)f\left(\delta'\right)d\delta'\\
&+\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}m\left(\lambda,\lambda'\right)\left\{\vartheta\left(\delta,a+q^e\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right],\lambda\right)-\vartheta\left(\delta,a,\lambda\right)\right.\\
&\left.+\vartheta\left(\delta',a'-q^e\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right],\lambda'\right)-\vartheta\left(\delta',a',\lambda'\right)\right\}\phi\left(\delta',a',\lambda'\right)\\
&\qquad d\delta'da'd\lambda'
\end{aligned}
$$

s.t.

$$
\vartheta_2\left(\delta,a+q^e\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right],\lambda\right)=\vartheta_2\left(\delta',a'-q^e\left[\left(\delta,a,\lambda\right),\left(\delta',a',\lambda'\right)\right],\lambda'\right).
$$

Checking that the planner's optimality conditions do not coincide with the equilibrium conditions is easy. More specifically, the comparison with Equation (3.13) reveals that the planner's optimality conditions and the equilibrium conditions would be identical if there was not $1/2$ in front of the matching function in the equilibrium condition. This difference is because of a composition externality typical of *ex post* bargaining environments, as discussed by Afonso and Lagos (2015). An individual investor of current type (δ, a, λ) internalizes only half the surpluses that her trades create. As a result, she does not internalize fully the social benefit that arises from the fact that having her in the current state (δ, a, λ) increases the meeting intensity of all other investors with an investor of type (δ, a, λ).

The solution method for the planner's problem is exactly the same as the solution method I used for equilibrium. In the end, the difference between the planner's solution and the equilibrium solution boils down to the use of a different endogenous inventory aversion. The inventory aversion that the benevolent social planner would assign to investors with λ solves the functional equation (4.7). The quantities chosen by the planner are given by (4.8).

Given the socially optimal trade quantities described above, the distribution of inventories solves the following system of Fourier transforms:

$$0 = -\left(\alpha + m\left(\lambda, \Lambda\right)\right) \widehat{g}^{e}_{\delta,\lambda}\left(z\right) + \alpha \int_{\delta_L}^{\delta_H} e^{-i2\pi\left(\delta' - \delta\right)C^{e}\left(\lambda\right)z} \widehat{g}^{e}_{\delta',\lambda}\left(z\right) f\left(\delta'\right) d\delta' + \int_0^M \int_{\delta_L}^{\delta_H} m\left(\lambda, \lambda'\right) \widehat{g}^{e}_{\delta,\lambda}\left(\frac{z}{1 + \frac{\widetilde{r}^{e}(\lambda')}{\widetilde{r}^{e}(\lambda)}}\right) \widehat{g}^{e}_{\delta',\lambda'}\left(\frac{z}{1 + \frac{\widetilde{r}^{e}(\lambda')}{\widetilde{r}^{e}(\lambda)}}\right) f\left(\delta'\right) \psi\left(\lambda'\right) d\delta' d\lambda' \tag{F.6}$$

for all $\lambda \in [0, M]$, $\delta \in [\delta_L, \delta_H]$ and for all $z \in \mathbb{R}$;

$$\widehat{g}^{e}_{\delta,\lambda}(0) = 1$$

for all $\lambda \in [0, M]$ and $\delta \in [\delta_L, \delta_H]$; and

$$\int_0^M \int_{\delta_L}^{\delta_H} \left(\widehat{g}^{e}_{\delta,\lambda}\right)'(0) f\left(\delta\right) \psi\left(\lambda\right) d\delta d\lambda = 0,$$

where

$$C^{e}\left(\lambda\right) \equiv \frac{1}{\kappa} \frac{\widetilde{r}^{e}\left(\lambda\right)}{\widetilde{r}^{e}\left(\lambda\right) + \alpha}.$$

So far, I have shown that the distortion of investors' decisions on the intensive margin leads to too cautious a trading behavior relative to the constrained efficient trading behavior. Next, I show how trade-size dependent transaction taxes/subsidies help eliminate this distortion. Suppose trading q units of the asset incurs a tax payment of $\tau_1(\lambda)(2aq + q^2)/2 + \tau_2(\lambda)\left(\delta - \overline{\delta}\right) q$ on the investor of type (δ, a, λ).[5] On the regulators' side, implementing such a policy in practice would require measuring the transaction frequencies of market participants and monitoring their risk exposures and asset positions. The recently implemented section of the Dodd-Frank Act, often referred to as "the Volcker Rule," which disallows proprietary trading by banks and their affiliates, also requires a similar

[5] Financial transaction taxes that are quadratic in trade size are also used in centralized market models, such as Subrahmanyam (1998) and Dow and Rahi (2000). The benefit of this specification is that it does not generate inaction regions in CARA-normal environments, and hence, allows for analytical and interior solution for trading rules.

level of monitoring. Some proprietary-trading forms are exempted from the Volcker Rule, such as those related to market making or hedging. Thus, regulators must monitor banks' positions and trading behavior and calculate certain metrics like transaction frequency or taste to determine proprietary trading unrelated to hedging or market making.

The bargaining problem of investors in the OTC market equilibrium with taxes will be

$$\begin{aligned}
&\{q\left[(\delta,a,\lambda),(\delta',a',\lambda')\right], P\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]\} \\
&= \arg\max_{q,P} \left[J(\delta,a+q,\lambda) - J(\delta,a,\lambda) - Pq - \frac{1}{2}\tau_1(\lambda)(2aq+q^2)\right. \\
&\left. - \tau_2(\lambda)\left(\delta-\overline{\delta}\right)q\right]^{\frac{1}{2}} \left[J(\delta',a'-q,\lambda') - J(\delta',a',\lambda') + Pq\right. \\
&\qquad\qquad \left. - \frac{1}{2}\tau_1(\lambda')(-2a'q+q^2) + \tau_2(\lambda')\left(\delta'-\overline{\delta}\right)q\right]^{\frac{1}{2}},
\end{aligned}$$

s.t.

$$\begin{aligned}
&J(\delta,a+q,\lambda) - J(\delta,a,\lambda) - Pq - \frac{1}{2}\tau_1(\lambda)(2aq+q^2) - \tau_2(\lambda)\left(\delta-\overline{\delta}\right)q \geq 0, \\
&J(\delta',a'-q,\lambda') - J(\delta',a',\lambda') + Pq - \frac{1}{2}\tau_1(\lambda')(-2a'q+q^2) \\
&\qquad\qquad + \tau_2(\lambda')\left(\delta'-\overline{\delta}\right)q \geq 0.
\end{aligned}$$

The first-order necessary and sufficient conditions and the Kuhn-Tucker conditions imply that the trade size, $q\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]$, maximizes the joint surplus net of total transaction tax; and the transaction price, $P\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]$, is set so that the maximized surplus net of total transaction tax is split equally between the bargaining parties; i.e., $q\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]$ and $P\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]$ solve the system

$$\begin{aligned}
&J_2(\delta,a+q,\lambda) - \tau_2(\lambda)\left(\delta-\overline{\delta}\right) - \tau_1(\lambda)a \\
&\quad = J_2(\delta',a'-q,\lambda') - \tau_2(\lambda')\left(\delta'-\overline{\delta}\right) - \tau_1(\lambda')a' + \left[\tau_1(\lambda)+\tau_1(\lambda')\right]q
\end{aligned}$$

$$\begin{aligned}
P &= \frac{J(\delta,a+q,\lambda) - J(\delta,a,\lambda) - \left(J(\delta',a'-q,\lambda') - J(\delta',a',\lambda')\right)}{2q} \\
&- \frac{1}{2}\left[\tau_2(\lambda)\left(\delta-\overline{\delta}\right) + \tau_2(\lambda')\left(\delta'-\overline{\delta}\right) + \tau_1(\lambda)a + \tau_1(\lambda')a'\right] \\
&\qquad\qquad - \frac{1}{4}\left[\tau_1(\lambda)-\tau_1(\lambda')\right]q.
\end{aligned}$$

Using this result, the HJB equation of investors becomes

$$
\begin{aligned}
rJ(\delta,a,\lambda) &= u(\delta,a)+T+\alpha\int_{\delta_L}^{\delta_H}[J(\delta',a,\lambda)-J(\delta,a,\lambda)]f(\delta')d\delta' \\
&+\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m(\lambda,\lambda')\frac{1}{2}\Big[\max_q\{J(\delta,a+q,\lambda)-J(\delta,a,\lambda)+J(\delta',a'-q,\lambda') \\
&-J(\delta',a',\lambda')-\left[\tau_2(\lambda)\left(\delta-\overline{\delta}\right)-\tau_2(\lambda')\left(\delta'-\overline{\delta}\right)+\tau_1(\lambda)a-\tau_1(\lambda')a'\right]q \\
&\qquad -\frac{1}{2}\left[\tau_1(\lambda)+\tau_1(\lambda')\right]q^2\Big\}\Big]\,\Phi(d\delta',da',d\lambda'),
\end{aligned}
$$

where

$$
\begin{aligned}
T &= \int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m(\lambda,\lambda')\Big(\frac{\tau_1(\lambda)}{2}\{2aq\left[(\delta,a,\lambda),(\delta',a',\lambda')\right] \\
&+\left(q\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]\right)^2\Big\}+\tau_2(\lambda)\left(\delta-\overline{\delta}\right)q\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]\Big) \\
&\qquad \Phi(d\delta',da',d\lambda')\Phi(d\delta,da,d\lambda)
\end{aligned}
$$

is the flow transfer from the government to investors.

The solution method for this problem is exactly the same as the solution method I used for equilibrium without taxes. The trade quantities in the equilibrium with taxes turn out to be

$$
\begin{aligned}
q\left[(\delta,a,\lambda),(\delta',a',\lambda')\right] &= \left[\frac{\kappa+r\tau_1(\lambda)}{\widetilde{r}(\lambda)}+\frac{\kappa+r\tau_1(\lambda')}{\widetilde{r}(\lambda')}\right]^{-1} \\
&\Big[-\frac{\kappa-\left(\widetilde{r}(\lambda)-r\right)\tau_1(\lambda)}{\widetilde{r}(\lambda)}\theta(\delta,a,\lambda)+\frac{\kappa-\left(\widetilde{r}(\lambda')-r\right)\tau_1(\lambda')}{\widetilde{r}(\lambda')}\theta(\delta',a',\lambda') \\
&\qquad -\tau_1(\lambda)a-\tau_2(\lambda)\left(\delta-\overline{\delta}\right)+\tau_1(\lambda')a'+\tau_2(\lambda')\left(\delta'-\overline{\delta}\right)\Big],
\end{aligned}
\tag{F.7}
$$

where

$$
\theta(\delta,a,\lambda)=a-A-\frac{1-\left(\widetilde{r}(\lambda)-r\right)\tau_2(\lambda)}{\kappa_1-\left(\widetilde{r}(\lambda)-r\right)\tau_1(\lambda)}\frac{\widetilde{r}(\lambda)}{\widetilde{r}(\lambda)+\alpha}\left(\delta-\overline{\delta}\right)
\tag{F.8}
$$

and

$$
\widetilde{r}(\lambda)=r+\int_0^M\frac{1}{2}m(\lambda,\lambda')\,\frac{\frac{\kappa-\left(\widetilde{r}(\lambda)-r\right)\tau_1(\lambda)}{\widetilde{r}(\lambda)}}{\frac{\kappa+r\tau_1(\lambda)}{\widetilde{r}(\lambda)}+\frac{\kappa+r\tau_1(\lambda')}{\widetilde{r}(\lambda')}}\psi(\lambda')d\lambda'.
\tag{F.9}
$$

Given this equilibrium trading behavior under the presence of taxes, the optimal policy presented in Proposition 8 is to choose $\tau_1(\lambda)$ and $\tau_2(\lambda)$ so that the equilibrium trade quantities (F.7) coincide with the constrained efficient trade quantities (4.8).

The social inefficiency in the OTC market equilibrium manifests itself in two intensive margin effects. First, investors' marginal valuation is more sensitive to inventories than the socially efficient marginal valuations. Thus, controlling for inventories, investors trade more cautiously leading to a less dispersed asset position distribution than the socially efficient asset position distribution. Second, in the calculation of (excess) inventories, investors put less weight on their current taste, which leads to less dispersed inventories. The roles of $\tau_1(\lambda)$ and $\tau_2(\lambda)$ are essentially to correct these two distortions, respectively.

Proposition 8 shows that $\tau_1(\lambda)$ is negative. This means that it is a subsidy whenever an investor with holding a trades in a way that her post-trade asset position is more extreme than $|a|$. Similarly, it is a tax whenever the investor ends up with a post-trade position less extreme than $|a|$. In short, $\tau_1(\lambda)$ gives investors incentive to increase the dispersion of asset position distribution. Over the lifetime of an investor, these taxes and subsidies stemming from terms with $\tau_1(\lambda)$ net out to zero.

In a similar fashion to $\tau_1(\lambda)$, $\tau_2(\lambda)$ gives investors incentive to make their inventories more dispersed. In particular, $\tau_2(\lambda)$ encourages an investor to sell when she has a strong taste for holding ($\delta > \overline{\delta}$) and encourages her to buy when she has a weak taste for holding ($\delta < \overline{\delta}$). Over an investor's lifetime, these taxes and subsidies stemming from terms with $\tau_2(\lambda)$ net out to a payment from the investor to the government simply because investors receive idiosyncratic taste shocks over time. During normal times, liquidity provision behavior typically leads to a subsidy and mean reversion to target holding leads to a tax, and these cancel each other out. However, immediately following an idiosyncratic shock, it takes the investor some time to reach her new target position, and she pays taxes during these episodes.

Proof of Proposition 8

Using $\tau_1(\lambda)$ specified in the proposition, (F.9) becomes:

$$\widetilde{r}(\lambda) = r + \int_0^M \frac{1}{2} m(\lambda, \lambda') \frac{\frac{\kappa}{\widetilde{r}^e(\lambda)} + \frac{\kappa}{\widetilde{r}^e(\lambda)} \frac{\widetilde{r}^e(\lambda) - r}{\widetilde{r}^e(\lambda) + r}}{\frac{\kappa}{\widetilde{r}^e(\lambda)} + \frac{\kappa}{\widetilde{r}^e(\lambda')}} \psi(\lambda') \, d\lambda',$$

where $\widetilde{r}^e(\lambda)$ is the solution of the corresponding functional equation (4.7) for the planner. Using (4.7), one notices that

$$\widetilde{r}(\lambda) = \frac{[\widetilde{r}^e(\lambda)]^2 + r^2}{\widetilde{r}^e(\lambda) + r} \Leftrightarrow \widetilde{r}^e(\lambda) = \frac{\widetilde{r}(\lambda) + \sqrt{[\widetilde{r}(\lambda)]^2 + 4r(\widetilde{r}(\lambda) - r)}}{2}.$$

After noticing this and using $\tau_1(\lambda)$ and $\tau_2(\lambda)$ specified in the proposition, it follows from (4.8), (4.9), (F.7), and (F.8) that

$$q^e\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]=q\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]$$

and

$$\theta^e(\delta,a,\lambda)=\theta(\delta,a,\lambda),$$

which establishes that the specified tax scheme decentralizes the constrained efficient allocation.

Now define and calculate, $\tau(\lambda)$, the instantaneous average financial transaction tax collected from investors with speed type λ:

$$\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m(\lambda,\lambda')\left(\frac{\tau_1(\lambda)}{2}\left\{2aq^e\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]\right.\right.$$
$$\left.\left.+\left(q^e\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]\right)^2\right\}+\tau_2(\lambda)\left(\delta-\overline{\delta}\right)q^e\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]\right)$$
$$\Phi(d\delta',da',d\lambda')\,\Phi_\lambda(d\delta,da)\equiv\tau(\lambda).$$

The integrand has three terms: The first two are related to $\tau_1(\lambda)$ and the last one is related to $\tau_2(\lambda)$. Let us calculate these terms one by one. The first term is:

$$\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m(\lambda,\lambda')\frac{\tau_1(\lambda)}{2}2aq^e\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]$$
$$\Phi(d\delta',da',d\lambda')\,\Phi_\lambda(d\delta,da)$$
$$=\tau_1(\lambda)\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m(\lambda,\lambda')\frac{-\widetilde{r}^e(\lambda')\,\theta^e(\delta,a,\lambda)+\widetilde{r}^e(\lambda)\,\theta^e(\delta',a',\lambda')}{\widetilde{r}^e(\lambda)+\widetilde{r}^e(\lambda')}$$
$$a\Phi(d\delta',da',d\lambda')\,\Phi_\lambda(d\delta,da)$$
$$=-\tau_1(\lambda)\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m(\lambda,\lambda')\frac{\widetilde{r}^e(\lambda')}{\widetilde{r}^e(\lambda)+\widetilde{r}^e(\lambda')}a\theta^e(\delta,a,\lambda)\,\Phi_\lambda(d\delta,da)\psi(\lambda')d\lambda'$$
$$=-\tau_1(\lambda)\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m(\lambda,\lambda')\frac{\widetilde{r}^e(\lambda')}{\widetilde{r}^e(\lambda)+\widetilde{r}^e(\lambda')}\left[\theta^e(\delta,a,\lambda)+C^e(\lambda)\left(\delta-\overline{\delta}\right)\right]$$
$$\theta^e(\delta,a,\lambda)\,\Phi_\lambda(d\delta,da)\,\psi(\lambda')\,d\lambda'$$

$$= -\tau_1(\lambda) \int_0^M m(\lambda, \lambda') \frac{\widetilde{r}^e(\lambda')}{\widetilde{r}^e(\lambda) + \widetilde{r}^e(\lambda')} \psi(\lambda')\, d\lambda'$$

$$\int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} \left[\theta^e(\delta, a, \lambda) + C^e(\lambda)\left(\delta - \overline{\delta}\right)\right] \theta^e(\delta, a, \lambda)\, \Phi_\lambda(d\delta, da)$$

$$= -\tau_1(\lambda)\left(\widetilde{r}^e(\lambda) - r\right)\left\{var\left[\theta^e|\lambda\right] + C^e(\lambda)\, cov\left[\delta, \theta^e|\lambda\right]\right\}.$$

The second term is:

$$\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m(\lambda, \lambda') \frac{\tau_1(\lambda)}{2} \left(q^e\left[(\delta, a, \lambda), (\delta', a', \lambda')\right]\right)^2$$

$$\Phi(d\delta', da', d\lambda')\, \Phi_\lambda(d\delta, da)$$

$$= \frac{\tau_1(\lambda)}{2} \int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m(\lambda, \lambda')$$

$$\left(\frac{-\widetilde{r}^e(\lambda')\,\theta^e(\delta, a, \lambda) + \widetilde{r}^e(\lambda)\,\theta^e(\delta', a', \lambda')}{\widetilde{r}^e(\lambda) + \widetilde{r}^e(\lambda')}\right)^2 \Phi(d\delta', da', d\lambda')\, \Phi_\lambda(d\delta, da)$$

$$= \frac{\tau_1(\lambda)}{2} \int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m(\lambda, \lambda') \left[\left(\frac{-\widetilde{r}^e(\lambda')\,\theta^e(\delta, a, \lambda)}{\widetilde{r}^e(\lambda) + \widetilde{r}^e(\lambda')}\right)^2\right.$$

$$\left. + \left(\frac{\widetilde{r}^e(\lambda)\,\theta^e(\delta', a', \lambda')}{\widetilde{r}^e(\lambda) + \widetilde{r}^e(\lambda')}\right)^2\right] \Phi(d\delta', da', d\lambda')\, \Phi_\lambda(d\delta, da)$$

$$= \frac{\tau_1(\lambda)}{2} \left[var\left[\theta^e|\lambda\right] \int_0^M m(\lambda, \lambda') \left(\frac{\widetilde{r}^e(\lambda')}{\widetilde{r}^e(\lambda) + \widetilde{r}^e(\lambda')}\right)^2 \psi(\lambda')\, d\lambda'\right.$$

$$\left. + \int_0^M m(\lambda, \lambda') \left(\frac{\widetilde{r}^e(\lambda)}{\widetilde{r}^e(\lambda) + \widetilde{r}^e(\lambda')}\right)^2 var\left[\theta^e|\lambda'\right] \psi(\lambda')\, d\lambda'\right].$$

By taking the derivative of (F.6) twice and evaluating it at $z = 0$ in the same fashion as the proof of Proposition 3, I obtain

$$\left(m(\lambda,\Lambda)-\int_0^M m(\lambda,\lambda')\left(\frac{\widetilde{r}^e(\lambda)}{\widetilde{r}^e(\lambda)+\widetilde{r}^e(\lambda')}\right)^2\psi(\lambda')\,d\lambda'\right)var\left[\theta^e|\lambda\right]$$
$$=-2C^e(\lambda)\,cov\left[\delta,\theta^e|\lambda\right]$$
$$+\int_0^M m(\lambda,\lambda')\left(\frac{\widetilde{r}^e(\lambda)}{\widetilde{r}^e(\lambda)+\widetilde{r}^e(\lambda')}\right)^2 var\left[\theta^e|\lambda'\right]\psi(\lambda')\,d\lambda'.$$

Substituting this into the previous expression, the second term of $\tau(\lambda)$ becomes

$$\frac{\tau_1(\lambda)}{2}\left[m(\lambda,\Lambda)var\left[\theta^e|\lambda\right]+var\left[\theta^e|\lambda\right]\int_0^M m(\lambda,\lambda')\frac{\left[\widetilde{r}^e(\lambda')\right]^2-\left[\widetilde{r}^e(\lambda)\right]^2}{\left[\widetilde{r}^e(\lambda)+\widetilde{r}^e(\lambda')\right]^2}\psi(\lambda')d\lambda'\right.$$
$$\left.2\left(\widetilde{r}^e(\lambda)-r\right)C^e(\lambda)\,cov\left[\delta,\theta^e|\lambda\right]\right]$$
$$=\frac{\tau_1(\lambda)}{2}\left[m(\lambda,\Lambda)\,var\left[\theta^e|\lambda\right]+var\left[\theta^e|\lambda\right]\int_0^M m(\lambda,\lambda')\frac{\widetilde{r}^e(\lambda')-\widetilde{r}^e(\lambda)}{\widetilde{r}^e(\lambda)+\widetilde{r}^e(\lambda')}\psi(\lambda')\,d\lambda'\right.$$
$$\left.2\left(\widetilde{r}^e(\lambda)-r\right)C^e(\lambda)\,cov\left[\delta,\theta^e|\lambda\right]\right]$$
$$=\frac{\tau_1(\lambda)}{2}\left[m(\lambda,\Lambda)\,var\left[\theta^e|\lambda\right]-m(\lambda,\Lambda)\,var\left[\theta^e|\lambda\right]+2\left(\widetilde{r}^e(\lambda)-r\right)var\left[\theta^e|\lambda\right]\right.$$
$$\left.2\left(\widetilde{r}^e(\lambda)-r\right)C^e(\lambda)\,cov\left[\delta,\theta^e|\lambda\right]\right]$$
$$=\tau_1(\lambda)\left[\left(\widetilde{r}^e(\lambda)-r\right)var\left[\theta^e|\lambda\right]+\left(\widetilde{r}^e(\lambda)-r\right)C^e(\lambda)\,cov\left[\delta,\theta^e|\lambda\right]\right].$$

Now one sees that the first and second terms of $\tau(\lambda)$ cancel each other out. Thus, only the last term will contribute. The last term is:

$$\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m(\lambda,\lambda')\,\tau_2(\lambda)\left(\delta-\overline{\delta}\right)q^e\left[(\delta,a,\lambda),(\delta',a',\lambda')\right]$$
$$\Phi(d\delta',da',d\lambda')\,\Phi_\lambda(d\delta,da)$$
$$=\tau_2(\lambda)\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H}\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m(\lambda,\lambda')\left(\delta-\overline{\delta}\right)$$
$$\frac{-\widetilde{r}^e(\lambda')\,\theta^e(\delta,a,\lambda)+\widetilde{r}^e(\lambda)\,\theta^e(\delta',a',\lambda')}{\widetilde{r}^e(\lambda)+\widetilde{r}^e(\lambda')}\Phi(d\delta',da',d\lambda')\,\Phi_\lambda(d\delta,da)$$
$$=-\tau_2(\lambda)\int_0^M\int_{-\infty}^{\infty}\int_{\delta_L}^{\delta_H} m(\lambda,\lambda')\frac{\widetilde{r}^e(\lambda')}{\widetilde{r}^e(\lambda)+\widetilde{r}^e(\lambda')}\left(\delta-\overline{\delta}\right)\theta^e(\delta,a,\lambda)$$
$$\Phi_\lambda(d\delta,da)\,\psi(\lambda')\,d\lambda'$$

$$= -\tau_2(\lambda) \int_0^M \int_{-\infty}^{\infty} \int_{\delta_L}^{\delta_H} m(\lambda, \lambda') \frac{\widetilde{r}^e(\lambda')}{\widetilde{r}^e(\lambda) + \widetilde{r}^e(\lambda')} \psi(\lambda') \, d\lambda' cov[\delta, \theta^e|\lambda]$$

$$= -\tau_2(\lambda)\left(\widetilde{r}^e(\lambda) - r\right) cov[\delta, \theta^e|\lambda].$$

Again, taking the derivative of (F.6) and evaluating it at $z = 0$ in the same fashion as the proof of Proposition 3 leads to:

$$cov[\delta, \theta^e|\lambda] = \frac{\alpha}{\alpha + \widetilde{r}^e(\lambda) - r} \frac{1}{\kappa} \frac{\widetilde{r}^e(\lambda)}{\widetilde{r}^e(\lambda) + \alpha} var[\delta].$$

Hence,

$$\tau(\lambda) = -\tau_2(\lambda) \frac{\alpha\left(\widetilde{r}^e(\lambda) - r\right)}{\alpha + \widetilde{r}^e(\lambda) - r} \frac{1}{\kappa} \frac{\widetilde{r}^e(\lambda)}{\widetilde{r}^e(\lambda) + \alpha} var[\delta].$$

After using $\tau_2(\lambda)$ defined in the proposition, the derivation of $\tau(\lambda)$ is complete.

APPENDIX G: DETAILS OF THE NETWORK MODEL

Bilateral trades. Using (5.1) and after simplification, (5.3) becomes

$$(q_{ij}, P_{ij}) = \arg\max_{q,P} \left(1 - e^{-\gamma\left[u\left(\delta_i, a^1_{-ij} + q_{ij}\right) - u\left(\delta_i, a^1_{-ij}\right) - q_{ij}P_{ij}\right]}\right)^{\frac{1}{2}} \left(1 - e^{-\gamma\left[u\left(\delta_j, a^1_{-ji} - q_{ij}\right) - u\left(\delta_j, a^1_{-ji}\right) + q_{ij}P_{ij}\right]}\right)^{\frac{1}{2}}, \tag{G.1}$$

s.t.

$$1 - e^{-\gamma\left[u\left(\delta_i, a^1_{-ij} + q_{ij}\right) - u\left(\delta_i, a^1_{-ij}\right) - q_{ij}P_{ij}\right]} \geq 0,$$
$$1 - e^{-\gamma\left[u\left(\delta_j, a^1_{-ji} - q_{ij}\right) - u\left(\delta_j, a^1_{-ji}\right) + q_{ij}P_{ij}\right]} \geq 0,$$

where a^1_{-ij} is investor i's post-trade asset position if she decides not to trade with investor j.

The solution (q_{ij}, P_{ij}) of the constrained optimization problem (G.1) satisfies the system

$$u_2\left(\delta_i, a^1_{-ij} + q_{ij}\right) = u_2\left(\delta_j, a^1_{-ji} - q_{ij}\right) \tag{G.2a}$$

$$P_{ij} = \frac{u\left(\delta_i, a^1_{-ij} + q_{ij}\right) - u\left(\delta_i, a^1_{-ij}\right) - \left(u\left(\delta_j, a^1_{-ji} - q_{ij}\right) - u\left(\delta_j, a^1_{-ji}\right)\right)}{2q_{ij}}. \tag{G.2b}$$

Using the definition of utility (5.2), the solution is

$$q_{ij} = \frac{a^1_{-ji} - a^1_{-ij}}{2} - \frac{1}{\kappa}\frac{\rho_j - \rho_i}{2}, \tag{G.3a}$$

$$P_{ij} = \overline{\delta} - \kappa\left(\frac{a^1_{-ij} + a^1_{-ji}}{2} - \frac{1}{\kappa}\frac{\delta_i + \delta_j}{2}\right). \tag{G.3b}$$

Using $a^1_{-ij} = a^1_i - q_{ij}$, (G.3a) can be written as

$$q_{ij} = a^1_{-ji} - a^1_i + \frac{1}{\kappa}\left(\delta_i - \delta_j\right).$$

Summing q_{ik} over all counterparties, k, of investor i, except for one particular counterparty j, one obtains Equation (5.4).

Equilibrium threat points. Equation (5.5) gives us a^1_{-ij} as a function of a^0_i, δ_i, q_{ij}, and λ_i. The main reason why the initial endowment, a^0_i, is a determinant of a^1_{-ij} is the price impact. The presence of price impact due to bargaining makes the investor unload her initial endowment to her counterparties imperfectly. Naturally, a^0_i enters the equation positively because even if the investor does not trade with investor j, a higher initial endowment leads to higher asset position for her. Secondly, the taste type, δ_i, enters the equation positively because higher δ_i means strong taste for holding the asset, and hence, the investor expects to buy more.

Substituting (5.5) into (G.3a) and (G3.b), all equilibrium objects can be written as a function of initial endowment, taste type, and number of counterparties, which leads to Proposition 9.

Comparison with the search model. Comparing (5.7) with (3.23) implies that the reciprocal of the number of counterparties has the role of determining the weight of an investor's inventory in the trade quantity in both models. However, the number of counterparties enters linearly in the network model, while it enters with a concave transformation in the search model. This means that the marginal liquidity provision incentive from having access to one additional counterparty stays constant in the network model, while it is decreasing in the search model. This difference arises due to the static vs. dynamic nature of the two models. In the search model, the calculation of $\widetilde{r}(\lambda)$ takes into account the fact that a fast investor's post-trade inventory in her future trades will be dictated, to a large extent, by her counterparties' trading needs, which creates a secondary negative impact of λ on $\widetilde{r}(\lambda)$ leading to concavity. This effect is missing in the static network model because an investor conducts all her trades simultaneously so she coordinates directly all her trades as shown by Equation (5.5).

Finally, comparing (5.8) with (3.24) reveals that there is no "connectedness" premium in the network model. As is clear from (G.2a) and (G.2b), the bargaining parties contribute equally to the trade surplus and then split it equally by taking the threat points as given. Thus, the speed premium term of (3.24) that appears in the search model does not appear in (5.8) of the network model.

Proof of Proposition 9

(5.5) implies

$$a^1_{-ij} = \frac{1}{\lambda_i} a^0_i + \frac{\lambda_i - 1}{\lambda_i} \left[A - q_{ij} - \frac{1}{\kappa} \left(\overline{\delta} - \delta_i \right) \right] \tag{G.4}$$

and

$$a^1_{-ji} = \frac{1}{\lambda_j} a^0_j + \frac{\lambda_j - 1}{\lambda_j} \left[A + q_{ij} - \frac{1}{\kappa} \left(\overline{\delta} - \delta_j \right) \right]. \tag{G.5}$$

Substituting these to (G.3a) and rearranging,

$$q_{ij} = \frac{-\frac{a^0_i - A}{\lambda_i} + \frac{1}{\kappa} \frac{\delta_i - \overline{\delta}}{\lambda_i} + \frac{a^0_j - A}{\lambda_j} - \frac{1}{\kappa} \frac{\delta_j - \overline{\delta}}{\lambda_j}}{\frac{1}{\lambda_i} + \frac{1}{\lambda_j}}, \tag{G.6}$$

which is equal to (5.7).

Substituting (G.4), (G.5), and (G.6) to (G.3b) and rearranging,

$$\begin{aligned} P_{ij} &= -\kappa \frac{a^0_i + (\lambda_i - 1) A - \frac{1}{\kappa} \left(\delta_i + (\lambda_i - 1) \overline{\delta} \right) + a^0_j}{\lambda_i + \lambda j} \\ &\quad - \kappa \frac{(\lambda_j - 1) A - \frac{1}{\kappa} \left(\delta_j + (\lambda_j - 1) \overline{\delta} \right)}{\lambda_i + \lambda j} \\ &= \overline{\delta} - \kappa A - \kappa \frac{a^0_i - A - \frac{1}{\kappa} \left(\delta_i - \overline{\delta} \right) + a^0_j - A - \frac{1}{\kappa} \left(\delta_j - \overline{\delta} \right)}{\lambda_i + \lambda j}, \end{aligned}$$

which is equal to (5.8).

APPENDIX H: MICRO-FOUNDATIONS FOR THE QUADRATIC UTILITY FLOW

Assume that there are two assets. One asset is riskless and pays interest at an exogenously given rate r. This asset is traded in a continuous frictionless market. The other asset is risky, traded over the counter, and is in supply denoted by A. This asset pays a cumulative dividend:

$$dD_t = m_D dt + \sigma_D dB_t,$$

where B_t is a standard Brownian motion.

I borrow the specification of preferences and trading motives from Duffie, Gârleanu, and Pedersen (2007) and Gârleanu (2009). Investors are subjective expected utility maximizers with CARA felicity functions. Investors' coefficient of absolute risk aversion and time preference rate are denoted by γ and r respectively.

Investor i has cumulative income process η^i:

$$d\eta_t^i = m_\eta dt + \sigma_\eta dB_t^i,$$

where

$$dB_t^i = \rho_t^i dB_t + \sqrt{1 - \left(\rho_t^i\right)^2} dZ_t^i.$$

The standard Brownian motion Z_t^i is independent of B_t, and ρ_t^i captures the instantaneous correlation between the payoff of the risky asset and the income of investor i. This correlation is time-varying and heterogeneous across investors. Thus, this heterogeneity creates the gains from trade. In the context of different markets, this heterogeneity can be interpreted in different ways such as hedging demands or liquidity needs. In the case of a credit derivatives market, for example, the correlation captures the exposure to credit risk. If a bank's exposure to the credit risk of a certain bond or loan is high, the correlation between the bank's income and the payoff of the derivative written on that specific bond or loan will be negative, implying that the derivative provides hedging to the bank. Therefore, that bank will have a high valuation for the derivative. Another bank with a short position in the bond will have a positive correlation and, consequently, a low valuation for the derivative.

I assume that the correlation between an investor's income and the payoff of risky asset is itself stochastic. Stochastic processes that govern idiosyncratic shocks and trade are as described in Section 2.

Let $V(W, \rho, a, \lambda)$ be the maximum attainable continuation utility of investor of type (ρ, a, λ) with current wealth W. It satisfies

$$V(W, \rho, a, \lambda) = \sup_c \mathbb{E}_t \left[-\int_t^\infty e^{-r(s-t)} e^{-\gamma c_s} ds \mid W_t = W,\ \rho_t = \rho,\ a_t = a \right],$$

s.t.

$$dW_t = (rW_t - c_t)dt + a_{t-} dD_t + d\eta_t - P\left[(\rho_{t-}, a_{t-}, \lambda), (\rho_t', a_t', \lambda_t')\right] da_t$$

$$da_t = \begin{cases} q\left[(\rho_{t-}, a_{t-}, \lambda), (\rho_t', a_t', \lambda_t')\right] & \text{if } (\rho_t', a_t', \lambda_t') \text{ is contacted} \\ 0 & \text{if no contact,} \end{cases}$$

where

$$\begin{aligned}\text{(H.1)} \quad & \{q\left[(\rho, a, \lambda), (\rho', a', \lambda')\right], P\left[(\rho, a, \lambda), (\rho', a', \lambda')\right]\} \\ & = \arg\max_{q,P} \left[V(W - qP, \rho, a + q, \lambda) - V(W, \rho, a, \lambda)\right]^{\frac{1}{2}} \\ & \qquad \left[V(W' + qP, \rho', a' - q, \lambda') - V(W', \rho', a', \lambda')\right]^{\frac{1}{2}},\end{aligned}$$

s.t.

$$\begin{aligned} & V(W - qP, \rho, a + q, \lambda) \geq V(W, \rho, a, \lambda), \\ & V(W' + qP, \rho', a' - q, \lambda') \geq V(W', \rho', a', \lambda').\end{aligned}$$

Since investors have CARA preferences, terms of trade are independent of wealth levels as I will show later. To eliminate Ponzi-like schemes, I impose the transversality condition

$$\lim_{T\to\infty} e^{-r(T-t)}\mathbb{E}_t\left[e^{-r\gamma W_T}\right] = 0.$$

To derive the optimal rules, the technique of stochastic dynamic programming is used. Assuming sufficient differentiability and applying Ito's lemma for jump-diffusion processes, the investor's value function $V(W, \rho, a, \lambda)$ satisfies the HJB equation

$$\begin{aligned}\text{(H.2)} \quad 0 = & \sup_c \{-e^{-\gamma c} + V_W(W, \rho, a, \lambda)[rW - c + am_D + m_\eta] \\ & + \frac{1}{2}V_{WW}(W, \rho, a, \lambda)[\sigma_\eta^2 + 2\rho a \sigma_D \sigma_\eta + a^2\sigma_D^2] \\ & - rV(W, \rho, a, \lambda) + \alpha\int_{-1}^{1}[V(W, \rho', a, \lambda) - V(W, \rho, a, \lambda)]f(\rho')d\rho' \\ & + \int_0^M\int_{-\infty}^{\infty}\int_{-1}^{1} m(\lambda, \lambda') \\ & \{V\left(W - q\left[(\rho, a, \lambda), x'\right] P\left[(\rho, a, \lambda), x'\right], \rho, a + q\left[(\rho, a, \lambda), x'\right], \lambda\right) \\ & \qquad - V(W, \rho, a, \lambda)\}\,\Phi(d\rho', da', d\lambda')\},\end{aligned}$$

where $x' \equiv (\rho', a', \lambda')$.

Following Duffie et al. (2007), I guess that $V(W, \rho, a, \lambda)$ takes the form

$$V(W, \rho, a) = -e^{-r\gamma(W + J(\rho, a, \lambda) + \overline{J})}$$

for some function $J(\rho, a)$, where

$$\overline{J} = \frac{1}{r}\left(m_\eta + \frac{\log r}{\gamma} - \frac{1}{2}r\gamma\sigma_\eta^2\right)$$

is a constant. Replacing into (H.2), I find that the optimal consumption is

$$c = -\frac{\log r}{\gamma} + r(W + J(\rho, a, \lambda) + \overline{J}).$$

After plugging c back into (H.2) and dividing by $r\gamma V(W, \rho, a, \lambda)$, I find that (H.2) is satisfied iff

$$\text{(H.3)} \qquad rJ(\rho, a, \lambda) = am_D - \frac{1}{2} r\gamma \left(a^2\sigma_D^2 + 2\rho a \sigma_D \sigma_\eta\right)$$
$$+ \alpha \int_{-1}^{1} \frac{1 - e^{-r\gamma[J(\rho', a, \lambda) - J(\rho, a, \lambda)]}}{r\gamma} f(\rho') d\rho'$$
$$+ \int_0^M \int_{-\infty}^{\infty} \int_{-1}^{1} \frac{1 - e^{-r\gamma\left\{J(\rho, a+q\left[(\rho,a,\lambda),x'\right], \lambda) - J(\rho,a,\lambda) - q\left[(\rho,a,\lambda),x'\right] P\left[(\rho,a,\lambda),x'\right]\right\}}}{r\gamma}$$
$$m\left(\lambda, \lambda'\right) \Phi(d\rho', da', d\lambda').$$

Terms of individual trades, $q\left[(\rho, a, \lambda), (\rho', a', \lambda')\right]$ and $P\left[(\rho, a, \lambda), (\rho', a', \lambda')\right]$, are determined by a Nash bargaining game with the solution given by the optimization problem (H.1). Dividing by $V(W, \rho, a, \lambda)^{\frac{1}{2}} V(W', \rho', a', \lambda')^{\frac{1}{2}}$, (H.1) can be written as

$$\left\{q\left[(\rho, a, \lambda), (\rho', a', \lambda')\right], P\left[(\rho, a, \lambda), (\rho', a', \lambda')\right]\right\}$$
$$= \arg\max_{q,P} \left[1 - e^{-r\gamma[J(\rho, a+q, \lambda) - J(\rho, a, \lambda) - qP]}\right]^{\frac{1}{2}}$$
$$\left[1 - e^{-r\gamma[J(\rho', a'-q, \lambda') - J(\rho', a', \lambda') + qP]}\right]^{\frac{1}{2}},$$

s.t.

$$1 - e^{-r\gamma[J(\rho, a+q, \lambda) - J(\rho, a, \lambda) - qP]} \geq 0$$
$$1 - e^{-r\gamma[J(\rho', a'-q, \lambda') - J(\rho', a', \lambda') + qP]} \geq 0.$$

As can be seen, terms of trade are independent of wealth levels. Solving this problem is relatively straightforward: I set up the Lagrangian of this problem. Then using the first-order and Kuhn-Tucker conditions, trade size $q\left[(\rho, a, \lambda), (\rho', a', \lambda')\right]$ solves (3.10). And, transaction price $P\left[(\rho, a, \lambda), (\rho', a', \lambda')\right]$ is given by (3.12) if $J_2(\rho, a, \lambda) \neq J_2(\rho', a', \lambda')$; and $P = J_2(\rho, a, \lambda)$ if $J_2(\rho, a, \lambda) = J_2(\rho', a', \lambda')$. Substituting the transaction price into (H.3), I get

$$
\begin{aligned}
rJ(\rho, a, \lambda) &= am_D - \frac{1}{2} r\gamma \left(a^2 \sigma_D^2 + 2\rho a \sigma_D \sigma_\eta \right) \\
&+ \alpha \int_{-1}^{1} \frac{1 - e^{-r\gamma[J(\rho', a, \lambda) - J(\rho, a, \lambda')]}}{r\gamma} f(\rho') d\rho' \\
&+ \int_0^M \int_{-\infty}^{\infty} \int_{-1}^{1} m(\lambda, \lambda') \\
&\frac{1 - e^{-\frac{r\gamma}{2}\left\{ J(\rho, a + q\left[(\rho, a, \lambda), x'\right], \lambda) - J(\rho, a, \lambda) + J(\rho', a' - q\left[(\rho, a, \lambda), x'\right], \lambda') - J(x') \right\}}}{r\gamma} \\
&\qquad \Phi(d\rho', da', d\lambda'),
\end{aligned}
\tag{H.4}
$$

subject to (3.10).

Eq. (H.4) cannot be solved in closed form. Consequently, following Gârleanu (2009), I use the linearization $\frac{1-e^{-r\gamma x}}{r\gamma} \approx x$ that ignores terms of order higher than 1 in $[J(\rho', a, \lambda) - J(\rho, a, \lambda)]$. The same approximation is also used by Biais (1993), Duffie et al. (2007), Vayanos and Weill (2008), and Praz (2014). Economic meaning of this approximation is that I assume investors are risk averse towards diffusion risks while they are risk neutral towards jump risks. The assumption does not suppress the impact of risk aversion as investors' preferences feature the fundamental risk-return trade-off associated with asset holdings. It only linearizes the preferences of investors over jumps in the continuation values created by trade or idiosyncratic shocks. The approximation yields the following lemma.

LEMMA 8 *Fix parameters $\overline{\gamma}$, $\overline{\sigma}_D$ and $\overline{\sigma}_\eta$, and let $\sigma_D = \overline{\sigma}_D \sqrt{\overline{\gamma}/\gamma}$ and $\sigma_\eta = \overline{\sigma}_\eta \sqrt{\overline{\gamma}/\gamma}$. In any stationary equilibrium, investors' value functions solve the following HJB equation in the limit as γ goes to zero:*

$$
\begin{aligned}
rJ(\rho, a, \lambda) &= am_D - \frac{1}{2} r\overline{\gamma} \left(a^2 \overline{\sigma}_D^2 + 2\rho a \overline{\sigma}_D \overline{\sigma}_\eta \right) \\
&+ \alpha \int_{-1}^{1} [J(\rho', a, \lambda) - J(\rho, a, \lambda)] f(\rho') d\rho' \\
&+ \int_0^M \int_{-\infty}^{\infty} \int_{-1}^{1} \frac{1}{2} m(\lambda, \lambda') \left\{ J(\rho, a + q\left[(\rho, a, \lambda), (\rho', a', \lambda')\right], \lambda) - J(\rho, a, \lambda) \right. \\
&\left. + J(\rho', a' - q\left[(\rho, a, \lambda), (\rho', a', \lambda')\right], \lambda') - J(\rho', a', \lambda') \right\} \Phi(d\rho', da', d\lambda'),
\end{aligned}
$$

subject to (3.10).

Setting $\kappa \equiv r\overline{\gamma}\,\overline{\sigma}_D^2$ and $\delta \equiv m_D - r\overline{\gamma}\,\overline{\sigma}_D\overline{\sigma}_\eta\rho$, the problem is equivalent to the one with the reduced-form quadratic utility flow.

APPENDIX I: TWO-DIMENSIONAL *EX ANTE* HETEROGENEITY

In this appendix, I consider a generalization of the baseline OTC model to two-dimensional *ex ante* heterogeneity: speed type, λ, and risk aversion parameter, γ:

$$u(\rho, a, \gamma) \equiv m_D a - \frac{1}{2} r\gamma \left(\sigma_D^2 a^2 + 2\sigma_D\sigma_\eta\rho a\right).$$

Let $\Psi(\lambda, \gamma)$ denote the joint cdf of speed types and risk aversion levels on $[0, M] \times [\gamma_L, \gamma_H]$. Speed types and risk aversion levels are allowed to be correlated but they are distributed independently from the hedging need types and from all the stochastic processes in the model. Differently from the baseline model, I assume $A = 0$ and $\overline{\rho} = 0$. In the baseline model without risk aversion heterogeneity, the result $\mathbb{E}_\phi[a \mid \lambda] = A$ obtains for an arbitrary positive A and an arbitrary $\overline{\rho}$. In this extended version, investors with low risk aversion levels want to have higher exposure to the aggregate endowment of risk, $A + \frac{\sigma_\eta}{\sigma_D}\overline{\rho}$. Thus, the result $\mathbb{E}_\phi[a \mid \lambda, \gamma] = A$ and the resulting simplifications afforded by the quadratic utility obtain only when $A = 0$ and $\overline{\rho} = 0$ in the extended model.

The investors' generalized problem (the counterpart of Equation (3.13)) can be written as

$$\begin{aligned} rJ(\rho, a, \lambda, \gamma) &= u(\rho, a, \gamma) + \alpha \int_{-1}^{1} [J(\rho', a, \lambda, \gamma) - J(\rho, a, \lambda, \gamma)] f(\rho') d\rho' \\ &+ \int_{\gamma_L}^{\gamma_H}\int_0^M \int_{-\infty}^{\infty}\int_{-1}^{1} m(\lambda, \lambda') \frac{1}{2} \Big[\max_q \ \{J(\rho, a+q, \lambda, \gamma) - J(\rho, a, \lambda, \gamma) \\ &\qquad + J(\rho', a'-q, \lambda', \gamma') - J(\rho', a', \lambda', \gamma')\}\Big] \Phi(d\rho', da', d\lambda', d\gamma'). \end{aligned}$$

To find the marginal valuation, I differentiate this equation with respect to a, applying the envelope theorem:

$$\begin{aligned} & rJ_2(\rho, a, \lambda, \gamma) = u_2(\rho, a, \gamma) \\ & + \alpha \int_{-1}^{1} [J_2(\rho', a, \lambda, \gamma) - J_2(\rho, a, \lambda, \gamma)] f(\rho') d\rho' \\ & + \int_{\gamma_L}^{\gamma_H}\int_0^M \int_{-\infty}^{\infty}\int_{-1}^{1} \frac{1}{2} m(\lambda, \lambda') \{J_2(\rho, a + q[(\rho, a, \lambda, \gamma), (\rho', a', \lambda')], \lambda) \\ & \qquad\qquad - J_2(\rho, a, \lambda, \gamma)\} \Phi(d\rho', da', d\lambda', \gamma'), \end{aligned}$$

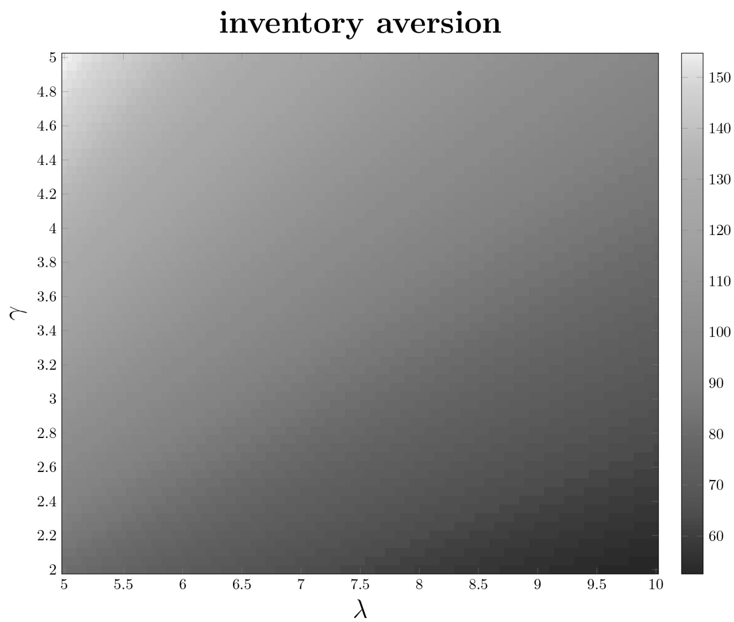

FIGURE 1.—Inventory aversion as a function of λ and γ, when $r = 0.05$, $\sigma_D = \sqrt{2000}$, $m(\lambda, \lambda') = 2\lambda\frac{\lambda'}{\Lambda}$, $\lambda \sim U[5, 10]$, $\gamma \sim U[2, 5]$, and λ and γ are independently distributed.

where

$$u_2(\rho, a, \gamma) = m_D - r\gamma\sigma_D^2 a - r\gamma\sigma_D\sigma_\eta\rho.$$

Following the exact same steps in the proof of Theorem 1 and Proposition 2, the equilibrium marginal valuation is

$$J_2(\rho, a, \lambda, \gamma) = \frac{m_D}{r} - \frac{\gamma\sigma_D^2}{\widetilde{r}(\lambda, \gamma)}\theta(\rho, a, \lambda, \gamma),$$

where

$$\theta(\rho, a, \lambda, \gamma) = a + \frac{\sigma_\eta}{\sigma_D}\frac{\widetilde{r}(\lambda, \gamma)}{\widetilde{r}(\lambda, \gamma) + \alpha}\rho$$

and $\widetilde{r}(\lambda, \gamma)$ solves the following generalized version of the functional equation (3.17):

$$\widetilde{r}(\lambda, \gamma) = r + \int_{\gamma_L}^{\gamma_H}\int_0^M \frac{1}{2} m(\lambda, \lambda') \frac{\frac{\gamma}{\widetilde{r}(\lambda,\gamma)}}{\frac{\gamma}{\widetilde{r}(\lambda,\gamma)} + \frac{\gamma'}{\widetilde{r}(\lambda',\gamma')}} \Psi(d\lambda', d\gamma'). \tag{I.1}$$

Here, the endogenous degree of inventory aversion of an investor is given by $\frac{\gamma\sigma_D^2}{\bar{r}(\lambda,\gamma)}$. In the baseline model without heterogeneity in risk aversion, λ was the only source of heterogeneity in investors' inventory aversion. Now, λ and γ jointly determine the inventory aversion.

Solving (I.1) numerically reveals that the inventory aversion is an increasing function of risk aversion and a decreasing function of speed type. Thus, Figure 1 shows that upward-sloping iso-inventory-aversion curves arise on the plane of risk aversion and trading speed because risk aversion and trading speed have opposite impact on the inventory aversion of an investor.

This generalization implies that if investors differ in their exogenous risk aversion levels as well as speed types, the main intermediaries are those with "low risk aversion and high speed type." Because these investors have the lowest endogenous inventory aversion, they have the comparative advantage in providing liquidity to others. As a result, investor centrality increases in the southeast direction of Figure 1.

APPENDIX J: THE CORPORATE BOND MARKET

TABLE I

DISTRIBUTION OF TRADE SIZES[a]

Sample	Obs.[b]	P1[b]	P10[b]	P50[b]	P90[b]	P99[b]	Mean[b]	SD[b]	Nrm. SD
2014									
A and above	2,979	1	5	31	1,220	10,000	631	2,825	4.47
Investment gr.	5,534	1	5	30	1,167	10,000	593	2,467	4.16
All bonds	8,941	1	5	43	1,410	10,000	599	2,523	4.21
2012–2014									
A and above	9,872	1	5	29	1,000	10,000	571	2,646	4.64
Investment gr.	18,323	1	5	28	1,000	10,000	525	2,293	4.37
All bonds	28,127	1	5	35	1,065	8,675	536	2,366	4.42
2005–2014									
A and above	32,939	1	5	25	1,000	10,000	549	3,233	5.89
Investment gr.	51,899	1	5	25	1,000	10,000	550	2,950	5.36
All bonds	75,246	1	5	25	1,325	10,000	587	3,463	5.90

[a]This table presents descriptive statistics for par value volume of transactions in the corporate bond market for the sample period from 2005 to 2014. "Sample" column specifies the subsample which statistics are based on. "P1," "P10," "P50," "P90," and "P99" show the 1st, 10th, 50th, 90th, and 99th percentile observation of the distribution, respectively. "Nrm. SD" (normalized standard deviation) is the ratio of sample standard deviation to sample mean.

[b]In thousands.

Trade Reporting and Compliance Engine (TRACE) was launched by the National Association of Securities Dealers (NASD) in 2002, by publicly reporting the transactions of approximately five hundred corporate bond issues of large and good credit entities at the beginning. The coverage expanded steadily over a few years, and by February 2005 it began disseminating 99% of all transactions in eligible corporate debt securities. I use enhanced TRACE database in this analysis, which includes trades that were not originally captured by standard TRACE database. I use the data filters proposed by Dick-Nielsen (2014) in cleaning enhanced TRACE data. This procedure eliminates potentially erroneous entries, reversals as well as canceled, corrected, and commissioned trades.

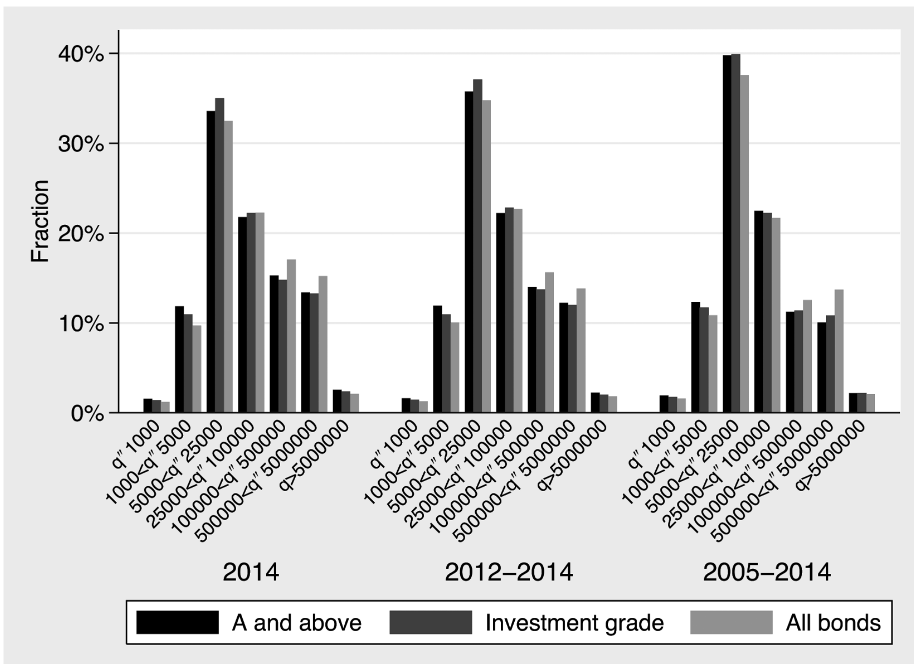

FIGURE 2.—This figure presents the distribution of corporate bond transactions across rating groups over different time periods. The sample includes all bond transactions obtained from TRACE. "q" represents the par value volume of the reported transaction. "2014," "2012–2014," and "2005–2014" indicate the three subsamples which distributions of trade sizes are presented. "A and above," "Investment grade," and "All bonds" show the trade size distributions of bonds with A and above credit rating, investment grade bonds, and all bonds, respectively.

REFERENCES

Afonso, G. and R. Lagos (2015): "Trade Dynamics in the Market for Federal Funds," *Econometrica*, 83, 263–313.

Biais, B. (1993): "Price Formation and Equilibrium Liquidity in Fragmented and Centralized Markets," *Journal of Finance*, 48, 157–185.

Brémaud, P. (1981): *Point Processes and Queues: Martingale Dynamics*, New York, NY: Springer-Verlag New York Inc.

Dick-Nielsen, J. (2014): "How to Clean Enhanced TRACE Data," Working paper.

Dow, J. and R. Rahi (2000): "Should Speculators be Taxed?" *Journal of Business*, 73, 89–107.

Duffie, D., N. Gârleanu, and L. H. Pedersen (2005): "Over-the-Counter Markets," *Econometrica*, 73, 1815–1847.

——— (2007): "Valuation in Over-the-Counter Markets," *Review of Financial Studies*, 20, 1865.

Gârleanu, N. (2009): "Portfolio Choice and Pricing in Illiquid Markets," *Journal of Economic Theory*, 144, 532–564.

Hutson, V., J. S. Pym, and M. J. Cloud (2005): *Applications of Functional Analysis and Operator Theory*, Amsterdam, the Netherlands: Elsevier B.V.

Lagos, R. and G. Rocheteau (2009): "Liquidity in Asset Markets with Search Frictions," *Econometrica*, 77, 403–426.

Martins-da Rocha, V. F. and Y. Vailakis (2010): "Existence and Uniqueness of a Fixed Point for Local Contractions," *Econometrica*, 78, 1127–1141.

Montes-de Oca, R. and E. Lemus-Rodríguez (2012): "An Unbounded Berge's Minimum Theorem with Applications to Discounted Markov Decision Processes," *Kybernetika*, 48, 268–286.

Praz, R. (2014): "Essays in Asset Pricing with Search Frictions," Ph.D. thesis, École Polytechnique Fédérale de Lausanne.

Protter, P. (2004): *Stochastic Integration and Differential Equations*, New York, NY: Springer.

Rincón-Zapatero, J. P. and C. Rodríguez-Palmero (2003): "Existence and Uniqueness of Solutions to the Bellman Equation in the Unbounded Case," *Econometrica*, 71, 1519–1555.

Stokey, N. L. and R. E. J. Lucas (1989): *Recursive Methods in Economic Dynamics*, Cambridge, MA: Harvard University Press.

Subrahmanyam, A. (1998): "Transaction Taxes and Financial Market Equilibrium," *Journal of Business*, 71, 81–103.

van Imhoff, E. (1982): *Optimal Economic Growth and Non-stable Population*, Berlin, Germany: Springer-Verlag.

Vayanos, D. and P.-O. Weill (2008): "A Search-based Theory of the On-the-Run Phenomenon," *Journal of Finance*, 63, 1361–1398.