University of Zurich[UZH]

Zurich Open Repository and Archive
University of Zurich
Main Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2015

Price distortions under coarse reasoning with frequent trade

Steiner, Jakub ; Stewart, Colin

Abstract: We study the effect of frequent trading opportunities and categorization on pricing of a risky asset. Frequent opportunities to trade can lead to large distortions in prices if some agents forecast future prices using a simplified model of the world that fails to distinguish between some states. In the limit as the period length vanishes, these distortions take a particular form: the price must be the same in any two states that a positive mass of agents categorize together. Price distortions therefore tend to be large when different agents categorize states in different ways, even if each individual's categorization is not very coarse.

DOI: https://doi.org/10.1016/j.jet.2015.07.011

Other titles: Price distortions in high-frequency markets

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-174511
Journal Article
Accepted Version

The following work is licensed under a Creative Commons: Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.

Originally published at:
Steiner, Jakub; Stewart, Colin (2015). Price distortions under coarse reasoning with frequent trade. Journal of Economic Theory, 159:574-595.
DOI: https://doi.org/10.1016/j.jet.2015.07.011

Supplementary Material for "Price Distortions in High-Frequency Markets"

Jakub Steiner*
Edinburgh University and CERGE-EI

Colin Stewart†
University of Toronto

October 27, 2014

1 Dynamic process

For convenience, we first restate the relevant elements of the model. A single asset pays a dividend that depends on a state $\omega(t)$ drawn from a finite set Ω. The state evolves according to an ergodic continuous-time stationary Markov process with transition rates $q(\omega, \omega')$. Trading occurs at discrete times $t = 0, \Delta, 2\Delta, \ldots$. We write ω_k for $\omega(k\Delta)$ and $q_\Delta(\omega, \omega')$ for the transition probabilities between trading periods. A constant per-unit flow dividend of $d(\omega_k)$ is paid from time $k\Delta$ to $(k+1)\Delta$.

A continuum of agents indexed by $i \in [0, 1]$ trades the asset in each period. Trading decisions are based on the current dividend and on agents' forecasts of the prices in the following period. Agents form these forecasts as follows. Each agent i categorizes states according to a partition Π^i of Ω that is fixed across all periods. For each state ω, let $\Pi(\omega)$ denote the element of the partition Π containing ω. In period k, agent i forms a forecast Q^i_{k+1} of the price in period $k+1$ according to

$$Q^i_{k+1} = \frac{\sum_{s<k-1:\omega_s \in \Pi^i(\omega_k)} p_{s+1}}{\sum_{s<k-1:\omega_s \in \Pi^i(\omega_k)} 1}$$

whenever the denominator is nonzero (otherwise take the forecast to be some arbitrary fixed number), where p_s denotes the market price in period s as described below. Thus the price forecast

*email: j-steiner@kellogg.northwestern.edu
†email: colinbstewart@gmail.com

Q^i_{k+1} is formed by averaging all prices that occurred in periods immediately following those in which the state was in the same category as the current one (according to Π^i).

Each agent i forms demand α^i_k in period k proportional to her net expected profit from holding the asset for one period:

$$\alpha^i_k = \left(1 - e^{-\Delta}\right) d(\omega_k) + e^{-\Delta} Q^i_{k+1} - p_k.$$

Assuming zero supply, the market clearing price is

$$p_k = \int_i p^i_k di, \tag{1}$$

where p^i_k is agent i's reservation price in period k, defined by

$$p^i_k = \left(1 - e^{-\Delta}\right) d(\omega_k) + e^{-\Delta} Q^i_{k+1}. \tag{2}$$

Let $\Pi_1, \ldots, \Pi_N$ denote those partitions belonging to a positive measure of agents, and denote by π_n the measure of agents using Π_n. Letting p^n_k denote the reservation price of each agent from group n, the market price p_k is

$$p_k = \sum_{n=1}^{N} \pi_n p^n_k.$$

2 Steady-state prices

Proposition 1 below shows that this learning process converges to steady-state prices $P : \Omega \longrightarrow \mathbb{R}$ that depend only on the current state. Steady-state prices turn out to be identical to rational expectations prices, not with respect to the true process, but with respect to a different process that reflects both the true process q_Δ and the categorizations used by agents.

Definition 1. Given any Δ, prices $P(\omega)$ are (steady-state) *rational expectations prices* with respect to a Markov process m on Ω and a dividend function d if

$$P(\omega) = (1 - e^{-\Delta}) d(\omega) + e^{-\Delta} E_{m(\omega,\omega')}[P(\omega')]$$

for every $\omega \in \Omega$.[1]

[1]Note that for any m and d, rational expectations prices exist and are unique.

Let ϕ denote the stationary distribution of states with respect to the true process q. For given initial prices and a given realization of the sequence of states $(\omega_s)_{s=0}^{k-1}$, let $p_k(\omega)$ denote the price in period k that would obtain if $\omega_k = \omega$. Define the *modified process* by

$$m_\Delta(\omega, \omega') = \sum_{n=1}^{N} \pi_n \sum_{\omega'' \in \Pi_n(\omega)} \phi\left(\omega'' \left| \Pi_n(\omega)\right.\right) q_\Delta(\omega'', \omega'). \tag{3}$$

Proposition 1. *For each Δ, the sequence $p_k(\omega)$ almost surely converges to the vector $P_\Delta(\omega)$ of rational expectations with respect to the modified process m_Δ and the dividend function d.*

Proposition 1 is a corollary of Proposition 2 below.

To understand the modified process m_Δ, first consider the case in which all agents distinguish all states, i.e. $\Pi^i(\omega) = \{\omega\}$ for every ω and i. In this this case, $m_\Delta = q_\Delta$, and hence the long-run prices are precisely the rational expectations prices with respect to the true process. To see why, consider the forecasting procedure. In period k, each agent uses data from previous periods $s < k-1$ in which the state was indistinguishable from the current state (according to her own categorization). For the finest categorization, these relevant periods are those s such that $\omega_s = \omega_k$. In the steady state, the agent's forecast is just the average of $P(\omega_{s+1})$ across the relevant periods s. In the long run, the forecast is equal to $\sum_{\omega'} q_\Delta(\omega_k, \omega')P(\omega')$, coinciding with the rational expectation of the price in the next period.

For general categorizations, a given agent's forecast is based on all previous periods s in which the state ω_s belonged to the current category $\Pi^i(\omega_k)$. In the long run, the average of $P(\omega_{s+1})$ for those values of s is equal to

$$\sum_{\omega'' \in \Pi^i(\omega)} \phi\left(\omega'' \left| \Pi^i(\omega)\right.\right) q_\Delta(\omega'', \omega')P(\omega'),$$

where the term $\phi\left(\omega'' \left| \Pi^i(\omega)\right.\right)$ captures the long-run frequency of state ω'' in the sample of relevant periods s. Taking the average across agents, the population-wide forecast is the expectation with respect to the modified process m_Δ in (3).[2]

The next section extends Proposition 1 in two directions. First, we extend the price forecasting

[2]The modified process is closely related to the coarse expectation formation in Eyster and Piccione (2013). Indeed, applying Proposition 1 to a homogenous population gives convergence to Eyster's and Piccione's stationary price function with respect to the agents' categorization.

rule to a general class of similarity-based rules in which agents forecast using data from similar past states. Unlike the categorization considered here, the weights assigned to different states may vary according to the perceived degree of similarity. Second, we allow for an arbitrary fraction of agents to form rational expectations knowing all parameters of the model, including other agents' forecasting procedures. In the long run, such agents have the same effect on prices as agents who categorize every state separately.

3 Proof and Generalizations

This section proves convergence of the above learning process, extends the result to a more general class of processes in which agents learn from similar past states, and shows that our results remain unchanged if we allow for some agents to form rational expectations. We start by describing learning by similarity, which includes categorization as a special case. We then consider an even more general class of processes that is sufficiently broad to allow for the inclusion of agents who form rational expectations about future states and other agents' behavior.

3.1 Learning by similarity

The categorization framework of Section 1 is a special case of a model in which agents learn prices based on past prices in states similar to the current one, but do not necessarily apply equal weight to all similar states. Proposition 1 extends to this more general case.

Each agent i is endowed with a symmetric similarity function $g_i : \Omega \times \Omega \longrightarrow \mathbb{R}_+$ determining the weight assigned to various states in forming forecasts of future prices. We assume that for each i and ω, there exists some ω' such that $g_i(\omega, \omega') \neq 0$. Given a history of states and prices up to period $k-1$, agent i's forecast in period k of the price in period $k+1$ is

$$Q^i_{k+1} = \frac{\sum_{s<k-1} g_i(\omega_k, \omega_s) p_{s+1}}{\sum_{s<k-1} g_i(\omega_k, \omega_s)}$$

whenever the denominator is nonzero, and some fixed constant otherwise. Thus the forecast is formed by averaging the one-period-ahead prices in all past states, weighted according to the degree of similarity to the current state. The categorization of Section 1 is a special case of this framework in which, for each i, g_i takes only the values 0 and 1.

For simplicity, we assume that only a finite number of different similarity functions are used by the agents. That is, there exists a finite partition of the population into groups of measures $\pi_1, \ldots, \pi_N$, and similarity functions $g_1, \ldots, g_N$ such that, for each n, every agent in n's group uses similarity function g_n. As before, each agent's action in each period is given by (2) and the market price in period k is the population-wide average action given by (1).

We show in the next subsection that Proposition 1 carries over directly to this setting except that the modified process m_Δ is defined more generally by

$$m_\Delta(\omega, \omega') = \sum_{n=1}^{N} \pi_n \frac{\sum_{\omega''} g_n(\omega, \omega'')\phi(\omega'')q_\Delta(\omega'', \omega')}{\sum_{\omega''} g_n(\omega, \omega'')\phi(\omega'')}. \tag{4}$$

Within each group, in the steady-state price forecasts, the weight given to each possible state ω' one period ahead is based on the likelihood of transitions to ω' from each state ω'' similar to the current state ω. The weight given to the transition from ω'' to ω' depends on the similarity between ω and ω'' together with the frequency $\phi(\omega'')$ with which state ω'' occurs. The aggregate distribution m_Δ is obtained by averaging the individual distributions across all agents. Note that, as before, agents should be interpreted as behaving, in the long-run, *as if* they believe (on average) that the state evolves according to m_Δ; agents do not literally hold these beliefs.

3.2 Proof of Proposition 1

The general learning process is as follows. The state space Ω and the true process q are as in the main text. Without loss of generality, let $\Delta = 1$. Let $\boldsymbol{\omega}^k = (\omega_s)_{s=0}^k$ denote the finite history of states up to period k, and $\mathbf{p}^k = (p_s)_{s=0}^{k-1}$ be the history of prices up to period $k-1$. We assume that all prices lie in a bounded interval $[\underline{p}, \overline{p}]$. The price p_k in period k is determined according to

$$p_k = (1-\rho)d(\omega_k) + \rho Q\left(\boldsymbol{\omega}^k, \mathbf{p}^k\right), \tag{5}$$

where $Q : \bigcup_k \left(\Omega^k \times [\underline{p}, \overline{p}]^{k-1}\right) \longrightarrow [\underline{p}, \overline{p}]$ can be interpreted as the average forecast of the price in period $k+1$ and $\rho = e^{-1}$ is the discount factor.

We assume that Q satisfies the following condition.

A1. There exists a continuous monotone function

$$\mathcal{E} : [\underline{p}, \overline{p}]^{\Omega} \longrightarrow [\underline{p}, \overline{p}]^{\Omega}$$

such that, for any $\underline{P}, \overline{P} \in [\underline{p}, \overline{p}]^{\Omega}$, any K, and any $\varepsilon > 0$, if

$$\Pr\left(p_k \in [\underline{P}, \overline{P}]\ \forall k > K\right) > 1 - \varepsilon, \tag{6}$$

then for any $\delta > 0$ there exists K' such that, for each ω,

$$\Pr\left(Q\left(\left(\boldsymbol{\omega}^k, \omega\right), \mathbf{p}^k\right) \in \left(\mathcal{E}(\underline{P})(\omega) - \delta, \mathcal{E}(\overline{P})(\omega) + \delta\right)\ \forall k > K'\right) > 1 - \varepsilon - \delta. \tag{7}$$

In the case of a homogeneous population using similarity function g, the learning process from Section 3.1 (and hence also the categorization-based learning from Section 1) is captured by

$$Q^{\text{sim}}\left(\boldsymbol{\omega}^k, \mathbf{p}^k\right) = \begin{cases} \frac{\sum_{s<k-1} g(\omega_k, \omega_s) p_{s+1}}{\sum_{s<k-1} g(\omega_k, \omega_s)} & \text{if } \sum_{s<k-1} g(\omega_k, \omega_s) > 0, \\ p_0 & \text{otherwise,} \end{cases}$$

where p_0 is arbitrary. For a heterogeneous population, Q is obtained by aggregating the values of Q^{sim} across groups (see Lemma 2).

Lemma 1. *For any similarity function g, Q^{sim} satisfies A1 with*

$$\mathcal{E}(P)(\omega) = \sum_{\omega'} m_{\Delta}(\omega, \omega') P(\omega'),$$

where m_{Δ} is the modified process in (4).

Proof. We prove only the upper bound; the proof for the lower bound is similar. Suppose that for some K, $\varepsilon > 0$, and $\overline{P}$, $\Pr\left(p_k \leq \overline{P}\ \forall k > K\right) > 1 - \varepsilon$. Given any $\delta > 0$ and $\gamma > 0$, by the Law of Large Numbers, there exists some $K' > K$ such that, with probability greater than $1-\delta$, for every pair (ω', ω'') and every $k > K'$, the fraction of periods $s < k$ such that $(\omega_s, \omega_{s+1}) = (\omega', \omega'')$ lies in $(\phi(\omega') q_{\Delta}(\omega', \omega'') - \gamma, \phi(\omega') q_{\Delta}(\omega', \omega'') + \gamma)$. Since the process q_{Δ} is ergodic, we can choose K' such that this property holds regardless of the history $\boldsymbol{\omega}^K$. Furthermore, for $K' > K/\gamma$, $p_s \leq \overline{P}(\omega_s)$

for a fraction of at least $1-\gamma$ periods $s \leq k$ with probability greater than $1-\varepsilon$, in which case the average of the prices p_{s+1} across those periods s such that $(\omega_s, \omega_{s+1}) = (\omega', \omega'')$ is at most $(1-\gamma)\overline{P}(\omega'') + \gamma\overline{p}$. Hence for $k > K'$, we have

$$Q^{\text{sim}}\left(\boldsymbol{\omega}^k, \mathbf{p}^k\right) \leq \frac{\sum_{\omega',\omega''} g(\omega_k, \omega'')(\phi(\omega'')q_\Delta(\omega'', \omega') + \gamma)((1-\gamma)\overline{P}(\omega') + \gamma\overline{p})}{\sum_{\omega''} g(\omega_k, \omega'')(\phi(\omega'') - \gamma)}$$

with probability greater than $1-\varepsilon-\delta$. Given $\delta > 0$, we can choose $\gamma > 0$ sufficiently small so that the right-hand side of the preceding inequality is less than $\mathcal{E}(\overline{P})(\omega_k) + \delta$, as needed. □

In the main text, agents differ in their forecasting procedures and the price is determined by the average of agents' forecasts. The following lemma indicates that A1 aggregates across heterogeneous groups.

Lemma 2. *Suppose that a fraction π_n of the population use prediction rule Q^n, with $\sum_{n=1}^N \pi_n = 1$. Suppose moreover that all rules Q^n satisfy A1 with functions $\mathcal{E}^n(P)$ respectively. Finally assume that price evolution is governed by* (5) *with prediction rule $Q = \sum \pi_n Q^n$. Then Q satisfies A1 with $\mathcal{E}(P) = \sum_n \pi_n \mathcal{E}^n(P)$.*

Proof. Using the property of A1 with $\pi_n\delta$ for each subpopulation and taking the maximum of the K' needed for each process gives the result. □

Proposition 1 is a special case of the following convergence result.

Proposition 2. *If Q satisfies A1, prices are determined according to* (5)*, and the mapping $(1-\rho)d + \rho\mathcal{E}$ is a contraction (with respect to some metric) then prices almost surely converge to the unique fixed point of $d + \rho\mathcal{E}$.*

In the case of learning by similarity, $(1-\rho)d + \rho\mathcal{E}$ is a contraction with respect to the sup norm, and we therefore obtain convergence to a unique price profile, proving Proposition 1.

Proof of Proposition 2. The mapping $(1-\rho)d+\rho\mathcal{E}$ has extreme fixed points $\underline{P}^*$, $\overline{P}^*$: for every fixed point P^*, we have $\underline{P}^* \leq P^* \leq \overline{P}^*$. This follows immediately from Tarski's Fixed Point Theorem since $[\underline{p}, \overline{p}]^\Omega$ is a complete lattice and $(1-\rho)d + \rho\mathcal{E}$ is continuous and monotone.

We will prove that for each ω, the set of cluster points of $(p_k(\omega))_k$ is almost surely contained in $[\underline{P}^*(\omega), \overline{P}^*(\omega)]$. The proposition follows immediately since the fixed point is unique when $(1-\rho)d + \rho\mathcal{E}$ is a contraction.

We prove only that the cluster points are almost surely at most $\overline{P}^*(\omega)$. The proof of the lower bound is similar.

Let $\overline{P}_0 = \overline{p}\mathbf{1}$, where $\mathbf{1}$ denotes the vector with a 1 in each component, and for $l \in \mathbb{N}_+$, let $\overline{P}_l = (1-\rho)d + \rho\mathcal{E}(\overline{P}_{l-1})$. Since $\overline{P}_l$ is nonincreasing in l, $\lim_l \overline{P}_l$ exists and is a fixed point of $(1-\rho)d + \rho\mathcal{E}$ (by continuity of $\mathcal{E}$).

Note that $p_k \leq \overline{P}_0(\omega_k)$ for each $k > 0$. Suppose for induction that, given any $\varepsilon > 0$, there exists K_l such that

$$\Pr\left(p_k < \overline{P}_l(\omega_k) + \varepsilon \text{ for all } k > K_l\right) > 1 - \varepsilon.$$

We will show that the same condition holds when each l is replaced with $l+1$.

For any $\delta > 0$, combining A1 with the inductive hypothesis, there exists some K_{l+1} such that

$$\Pr\left(Q\left(\boldsymbol{\omega}^k, \mathbf{p}^k\right) < \mathcal{E}(\overline{P}_l + \varepsilon\mathbf{1})(\omega_k) + \delta\, \forall k > K_{l+1}\right) > 1 - \varepsilon - \delta.$$

Substituting for $Q\left(\boldsymbol{\omega}^k, \mathbf{p}^k\right)$ using (5), we have

$$\Pr\left(p_k < (1-\rho)d(\omega_k) + \rho\mathcal{E}(\overline{P}_l + \varepsilon\mathbf{1})(\omega_k) + \delta\, \forall k > K_{l+1}\right) > 1 - \varepsilon - \delta.$$

Given any $\gamma > 0$, since $\mathcal{E}$ is continuous, there exist some $\varepsilon, \delta \in (0, \gamma)$ such that, for each ω, $\rho\mathcal{E}(\overline{P}_l + \varepsilon\mathbf{1})(\omega) + \delta < \rho\mathcal{E}(\overline{P}_l)(\omega) + \gamma$. Since ε and δ are arbitrary, we have that, for some K_{l+1},

$$\Pr\left(p_k < (1-\rho)d(\omega_k) + \rho\mathcal{E}(\overline{P}_l)(\omega_k) + \gamma\, \forall k > K_{l+1}\right) > 1 - \gamma.$$

Since $\overline{P}_{l+1} = (1-\rho)d + \rho\mathcal{E}(\overline{P}_l)$, this completes the proof of the inductive step. □

3.3 Presence of rational agents

We now consider a setting in which some agents form rational expectations. For simplicity, we assume that the population consists of two parts. A fraction π of agents are rational while the remaining $1-\pi$ are coarse thinkers who use a prediction rule Q^C satisfying A1. Rational agents know Q^C and the underlying Markov process, and form rational expectations of the forecasts formed by coarse thinkers in the next period. The rational agents' prediction rule Q^R satisfies

$$Q^R\left(\boldsymbol{\omega}^k, \mathbf{p}^k\right) = E\left[(1-\rho)d(\omega_{k+1}) + \rho\left((1-\pi)Q^C\left(\boldsymbol{\omega}^{k+1}, \mathbf{p}^{k+1}\right) + \pi Q^R\left(\boldsymbol{\omega}^{k+1}, \mathbf{p}^{k+1}\right)\right)\Big|\, \boldsymbol{\omega}^k\right]. \tag{8}$$

This equation implies that rational agents correctly predict prices given the history to date and the prediction rules used by other agents.

While the model in the main text does not include rational agents, it does allow for some agents to perfectly distinguish among states. In the present setting, these agents are not rational insofar as their price forecasts are based only on past data and do not explicitly account for other agents' forecasts. We show here that, in the long-run, the difference between these agents and rational agents is immaterial. Long-run prices are identical if we replace any share of agents using the finest categorization with agents who form rational expectations.

A2. For each $\omega \in \Omega$, $K \in \mathbb{N}$, and almost every $\boldsymbol{\omega} \in \Omega^{\mathbb{N}}$,

$$\lim_{k\to\infty}\left(Q^C\left(\left(\boldsymbol{\omega}^k, \omega\right), \mathbf{p}^k\right) - Q^C\left(\left(\boldsymbol{\omega}^{k+K}, \omega\right), \mathbf{p}^{k+K}\right)\right) = 0,$$

where, for each κ, $\boldsymbol{\omega}^\kappa$ denotes the projection of $\boldsymbol{\omega}$ onto its first κ components.

Roughly speaking, A2 says that data from a fixed finite number of recent periods eventually has little impact on forecasts once the total quantity of data is large. Note that A2 is satisfied by the similarity-based learning procedure of Section 3.1.

Proposition 3. *Suppose that a fraction π of the population form rational expectations, and the remaining $1-\pi$ use a prediction procedure Q^C satisfying A1 with bound $\mathcal{E}^C(P)$ and A2. Suppose further that the mapping $(1-\rho)d + \rho\mathcal{E}^C(P)$ is a contraction. Then the price vector $P(\omega)$ almost surely converges to the unique solution of*

$$P(\omega_k) = (1-\rho)d(\omega_k) + \rho\left(\pi E\left[P\left(\omega_{k+1}\right) \mid \omega_k\right] + (1-\pi)\mathcal{E}^C(P)(\omega_k)\right).$$

Lemma 3. *If Q^C satisfies A1 with bound $\mathcal{E}^C$ and A2 then Q^R satisfies A1 with bound*

$$\mathcal{E}^R(P)(\omega_k) = E\left[\sum_{l=1}^{\infty}(\pi\rho)^{l-1}(1-\rho)d(\omega_{k+l}) + (1-\pi)\rho\sum_{l=1}^{\infty}(\pi\rho)^{l-1}\mathcal{E}^C(P)(\omega_{k+l})\right|\omega_k\Bigg]. \tag{9}$$

Proof of Lemma 3. Iterating (8) gives

$$Q^R\left(\boldsymbol{\omega}^k, \mathbf{p}^k\right) = E\left[\sum_{l=1}^{\infty}(\pi\rho)^{l-1}(1-\rho)d(\omega_{k+l}) + (1-\pi)\rho\sum_{l=1}^{\infty}(\pi\rho)^{l-1}Q^C\left(\boldsymbol{\omega}^{k+l}, \mathbf{p}^{k+l}\right)\right|\boldsymbol{\omega}^k\Bigg].$$

We need to show that for any $\underline{P}, \overline{P} \in [\underline{p}, \overline{p}]^{\Omega}$, any K, and any $\varepsilon > 0$, if condition (6) holds, then for any $\delta > 0$ there exists K' such that (7) holds for Q^R. We prove only the upper bound; the proof of the lower bound is similar.

Accordingly, suppose that (6) holds for some $\varepsilon > 0$ and K. Fix $\delta > 0$. Since Q^C and $\mathcal{E}^C$ are bounded, there exists M such that, for every $\boldsymbol{\omega}^k$ and $\mathbf{p}^k$

$$Q^R\left(\boldsymbol{\omega}^k, \mathbf{p}^k\right) < E\left[\sum_{l=1}^{\infty}(\pi\rho)^{l-1}(1-\rho)d(\omega_{k+l}) + (1-\pi)\rho\left(\sum_{l=1}^{M}(\pi\rho)^{l-1}Q^C\left(\boldsymbol{\omega}^{k+l}, \mathbf{p}^{k+l}\right) + \sum_{l=M+1}^{\infty}(\pi\rho)^{l-1}\mathcal{E}^C(\overline{P})(\omega_{k+l})\right)\right|\boldsymbol{\omega}^k\Bigg] + \delta/3. \tag{10}$$

Since Q^C satisfies A1, there exists some K' such that, for each ω,

$$\Pr\left(Q^C\left(\left(\boldsymbol{\omega}^{k-1}, \omega\right), \mathbf{p}^k\right) < \mathcal{E}^C(\overline{P})(\omega) + \delta/3M\ \forall k > K'\right) > 1 - \varepsilon - \delta/2. \tag{11}$$

By A2, there exists some K'' such that, for each $l = 1, \ldots, M$,

$$\Pr\left(Q^C\left(\boldsymbol{\omega}^{k+l}, \mathbf{p}^{k+l}\right) < Q^C\left(\left(\boldsymbol{\omega}^{k-1}, \omega_{k+l}\right), \mathbf{p}^k\right) + \delta/3M\ \forall k > K'' \mid \boldsymbol{\omega}^k\right) > 1 - \delta/2M. \tag{12}$$

Combining (11) and (12) gives

$$\Pr\left(Q^C\left(\boldsymbol{\omega}^{k+l}, \mathbf{p}^{k+l}\right) < \mathcal{E}^C(\overline{P})(\omega_{k+l}) + 2\delta/3M\ \forall k > \max\{K', K''\},\ \forall l = 1, \ldots, M \mid \boldsymbol{\omega}^k\right) > 1 - \varepsilon - \delta.$$

Combining the last inequality with (10) gives

$$\Pr\left(Q^R\left(\boldsymbol{\omega}^k, \mathbf{p}^k\right) < \mathcal{E}^R(\overline{P})(\omega_k) + \delta\, \forall k > \max\{K', K''\}\right) > 1 - \varepsilon - \delta,$$

as needed. □

Proof of Proposition 3. Combining Lemma 2, Lemma 3, and Proposition 2, the cluster points lie between the extremal solutions to

$$P = (1-\rho)d + \rho\left(\pi\mathcal{E}^R(P) + (1-\pi)\mathcal{E}^C(P)\right).$$

Substituting for $\mathcal{E}^R(P)$ using (9) leads to

$$\begin{aligned}
P(\omega_k) &= (1-\rho)d(\omega_k) + \rho\pi E\left[\sum_{l=1}^{\infty}(\pi\rho)^{l-1}(1-\rho)d(\omega_{k+l}) + (1-\pi)\rho\sum_{l=1}^{\infty}(\pi\rho)^{l-1}\mathcal{E}^C(P)(\omega_{k+l})\middle|\,\omega_k\right] \\
&\quad + \rho(1-\pi)\mathcal{E}^C(P)(\omega_k) \\
&= (1-\rho)d(\omega_k) + \rho\pi E\left[E\left[\sum_{l=1}^{\infty}(\pi\rho)^{l-1}(1-\rho)d(\omega_{k+l}) + (1-\pi)\rho\sum_{l=1}^{\infty}(\pi\rho)^{l-1}\mathcal{E}^C(P)(\omega_{k+l})\middle|\,\omega_{k+1}\right]\middle|\,\omega_k\right] \\
&\quad + \rho(1-\pi)\mathcal{E}^C(P)(\omega_k) \\
&= (1-\rho)d(\omega_k) + \rho\left(\pi E[P(\omega_{k+1})|\omega_k] + (1-\pi)\mathcal{E}^C(P)(\omega_k)\right),
\end{aligned}$$

where the second last equality follows from the Law of Iterated Expectations, and the final equality uses the first equality with ω_{k+1} in place of ω_k. □

References

Eyster, E. and M. Piccione (2013). An approach to asset pricing under incomplete and diverse perceptions. *Econometrica 81*(4), 1483–1506.