Trading Volume in General Equilibrium with Complete Markets

Eric M. Aldrich*

Department of Economics
University of California, Santa Cruz

January 28, 2013

Abstract

This paper investigates asset trade in a general-equilibrium complete-markets endowment economy with heterogeneous agents. It shows that standard no-trade results cease to hold when agents have heterogeneous beliefs and that substantial trade volume is generated, even in the presence of a spanning set of assets. Further, trade volume and price movements have a positive relationship in the model, as is well documented in the empirical literature. This paper also develops a computational algorithm for solving finite-period heterogeneous-beliefs economies and demonstrates how the problem is well suited for large-scale parallel computing methods, such as GPU computing.

Keywords: Trading Volume, Asset Pricing, Heterogeneous Beliefs, General Equilibrium.

JEL Classification: C60, C68, D5, D80, E21, G12.

*Email: ealdrich@gmail.com; Phone: 919-638-6216. The views expressed herein are those of the author and do not necessarily represent those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.
1 Introduction

It is well understood that in standard complete-markets endowment economies, agents do not engage in trade. Lucas (1978) is the canonical development of this fact. Judd et al. (2003) shows that similar results hold even when agents are heterogeneous in their endowments and preferences. In particular, Judd et al. (2003) shows that if the spanning set of assets is infinitely-lived, agents will (potentially) engage in trade at time zero, but never again. When the spanning set includes a mix of finite-maturity assets, agents may engage in trade at each time period, but only in such a way that allows them to maintain constant portfolios through time. The results of Judd et al. (2003) depend crucially on the Markovian properties of their endowment economy, which leads to time-homogeneous consumption policy rules.

Numerous empirical studies have shown trading volume to be important for understanding market structure, return volatility, and the conditional distribution of asset prices. Karpoff (1987), Gallant et al. (1992) and Tauchen et al. (1996) are only a few examples from a very large literature. Few theoretical models exist which explain volume in equilibrium, and those that do typically resort to a partial equilibrium approach or incorporate some form of market incompleteness or asymmetric information. This paper falls in the category of the macroeconomic asset pricing literature, which seeks to explain features of financial markets within a macroeconomic framework. In particular, it shows how heterogeneity of beliefs leads to nontrivial asset exchange among agents in general equilibrium and with complete markets. The same is not generally true of other forms of agent heterogeneity.

The results of this paper show that diversity of beliefs creates a motive for agents to bet against each other on paths of the economy that they view as relatively more likely. As conceptualized by Simsek (2011), this speculative phenomenon is reflected in increased variability of agents’ net worths. The desire to increase consumption on certain paths and decrease it on others introduces a history dependence and lack of time homogeneity in consumption plans which induces nontrivial trade at dates beyond the initial period, even when the economy is Markovian. The upshot is that asset exchange is fundamentally related to agents’ diverse views. Consistent with empirical studies, the results herein also show that volume is positively related to the degree of belief dispersion in addition to the proportion
of agents that hold divergent beliefs. All of this is achieved without short-sales constraints and in time periods where no unanticipated information events are realized.

Agents’ desire to speculate also has implications for asset prices. In particular, introducing agents with divergent views causes prices to rise in states that those divergent agents view as relatively more likely. This reflects the speculative phenomenon suggested by Harrison and Kreps (1978): agents are willing to purchase assets at a higher price than they intrinsically value them due to the presence of another agent class with a differing valuation. Importantly, the relationship of heterogeneous beliefs to both asset exchange and price movements, leads trade volume within the model to be positively related to absolute price changes, a feature which has been well documented in the empirical literature.

The importance of heterogeneous beliefs in generating trading volume is illustrated in an empirical example which is congruent with historical data. Using quarterly post-war consumption data to estimate parameters of the endowment process and beliefs, this paper shows that sizeable trading volume can exist in general equilibrium with complete markets, while simultaneously matching broad features of the macroeconomy. Even a mild disturbance in beliefs can generate substantial volume and price movements.

The model in this paper does not allow for learning, although such a feature would be a desirable extension for future work. While the fundamental trade result does not depend on a finite-horizon economy, it is well understood that under the paradigm considered herein, only one agent type (the one whose beliefs are most correct) will eventually survive in the market. This paper does not concern itself with questions of survival. The truth is that questions of survival often take a very long time to resolve, and in the finite-period application of this model, survival has a negligible impact over the course of 20 to 30 years.

This paper also develops a computational algorithm for solving a finite-period version of the model. The complete-markets structure allows for optimal consumption choices and prices at each date and state to be independent of all but the first state. The result is that the competitive equilibrium can be solved with a procedure which iterates on the initial consumption choices, solving a large number of independent problems at each future state. The independence of the problems at each subsequent state suggests that a parallel computing methodology can be used to solve for the competitive equilibrium quickly and efficiently. In
particular, this is an example of an economic problem that is well suited for GPU computing, as discussed in Aldrich et al. (2011).

2 Model

We consider a basic endowment economy with $I$ types of agents and $S$ aggregate states of nature each period. Time is discrete and indexed by $t \in \mathcal{N}_0 = \{0, 1, 2, \ldots\}$. The aggregate state at time $t$ is $s_t \in \mathcal{S} = \{1, \ldots, S\}$ and we let $s^t = (s_0, s_1, \ldots, s_t)$ denote the history of aggregate states. Agents’ types are indexed by $i \in \mathcal{I} = \{1, 2, \ldots, I\}$ and $\mu^i(s^t)$ denotes the proportion of the population consisting of type $i$ agents in state $s^t$. The total population has unit mass, which dictates $\sum_{i=1}^{I} \mu^i(s^t) = 1$, for all $s^t \in \mathcal{S}^t$.

There is a single consumption good and a tree paying a dividend of $d(s^t)$ units of the consumption good in each state $s^t \in \mathcal{S}^t$, which cannot be transferred between time periods. By default, each agent is entitled to $d(s^t)$ units of consumption in state $s^t$, resulting in an endowment of $\mu^i(s^t)d(s^t)$ for cohort $i$ and an aggregate endowment of $\sum_{i=1}^{I} \mu^i(s^t)d(s^t) = d(s^t)$. Agents have preferences for consumption encapsulated in period utility $u_i(c)$, which is type specific and which satisfies the usual conditions of strict monotonicity, strict concavity, twice continuous differentiability and $\lim_{c \to 0} u_i'(c) = \infty$.

The aggregate state follows an $S$-state Markov process: the probability of history $s^t$ is $\pi(s^t) = \pi(s_t|s_{t-1}) \cdots \pi(s_1|s_0) \pi(s_0)$, where $s_0 \in \mathcal{S}$ is known and hence $\pi(s_0) = 1$. Aside from preference and endowment heterogeneity, encapsulated in $u_i(c)$ and $\mu^i(s^t)d(s^t)$, respectively, we allow agent types to have heterogeneous discount factors, $\beta_i$, and heterogeneous beliefs about transition probabilities, $\pi^i(s^t) = \pi^i(s_t|s_{t-1}) \cdots \pi^i(s_1|s_0)$. Discount factor heterogeneity is frequently considered a form of preference heterogeneity; for the purposes of this paper, it will be important to distinguish between heterogeneity of period utility functions (referred to as preference heterogeneity) and that of discount factors.

Markets are complete and agent types can deviate from their endowments by purchasing state-contingent consumption, $c_i^t(s^t)$. As discussed below, the results of the paper are unchanged for general asset markets, allowing agents access to assets with varying maturities and payoff structures, so long as markets are complete. The results, however, are easiest
to understand within the framework of state-contingent consumption purchases. We denote the time zero price of consumption in state $s^t$ as $q^0(s^t)$. The resulting optimization problem for an individual agent of type $i$ is

$$U_i(c^i) = \max_{c^i} \left\{ u(c^i(s_0)) + \sum_{t=1}^{\infty} \beta_t^i \sum_{s^t} u(c^i(s^t)) \pi^i(s^t) \right\}$$

(1a)

subject to

$$c^i(s_0) + \sum_{t=1}^{\infty} \sum_{s^t} q^0(s^t)c^i(s^t) \leq d(s_0) + \sum_{t=1}^{T} \sum_{s^t} q^0(s^t)d(s^t),$$

(1b)

where $c^i = (c(s_0), c(s^1), \ldots)$ and where $q^0(s_0) = 1$.

A competitive equilibrium for this economy is a collection of consumption plans $\{\bar{c}^i\}_{i=1}^I$ and prices $\{\{q^0(s^t)\}_{s^t \in S^t}\}_{t=0}^\infty$ such that

1. System (1) is solved.

2. The aggregate resource constraint

$$\sum_{i=1}^{I} \mu^i(s^t)\bar{c}^i(s^t) = \sum_{i=1}^{I} \mu^i(s^t)d(s^t)$$

holds for all $s^t \in S^t$ and $t \geq 0$.

### 2.1 Asset Markets

Consider $S$ long-lived securities, indexed by $j \in S$, with asset $j$ paying the aggregate dividend $d(s^{t-1}, j)$ if $s_t = j$ and zero otherwise, for all $t > 0$. As an alternative to state-contingent consumption purchases, agents can achieve their optimal consumption plans if given access to these $S$ securities. At each period, the budget constraint for agent $i$ is

$$c^i(s^t) + \sum_{j=1}^{S} q^0(s^t, j)\theta^i_j(s^t) \leq d(s^t)\theta^i_i(s^{t-1}) + \sum_{j=1}^{S} q^0(s^t, j)\theta^i_j(s^{t-1}).$$

(3)

where $q^0(s^{t-1}, j)$ denotes $q^0(s^t)$ for the history $s^t = (s_0, \ldots, s_{t-1}, s)$, $s \in S$, and where $\theta^i_{s_0}(s^{-1})$ represents the endowment of agent $i$ (in shares) at $t = 0$. Equation (3) states that agents enter period $t$ with shares $\{\theta^i_j(s^{t-1})\}_{j=1}^{S}$, only one of which pays the aggregate dividend,
\( \theta_i^j(s^{t-1}) \). After receiving a portion of the aggregate dividend and liquidating all shares for their market prices, \( q_0^i(s^t, j) \), agent \( i \) has total resources that are equivalent to the right-hand side of Equation (3). She then allocates her resources according to the left-hand side of (3). The volume of trade for each asset and each agent in each period amounts to

\[
TV_j^i(s^t) = \left| \theta_j^i(s^t) - \theta_j^i(s^{t-1}) \right| \sum_{i=1}^I \mu^i(s^t) \theta_j^i(s^{t-1}),
\]

where the denominator in (4) is always unity. The total volume of trade for each asset in each period is then

\[
TV_j(s^t) = \sum_{i=1}^I \mu^i(s^t) TV_j^i(s^t).
\]

### 3 Trade in General Equilibrium

The fundamental result of this paper is that belief heterogeneity can generate nontrivial exchange among agent types, even when the entire structure of the economy is Markovian. The same is not true of other forms of heterogeneity (discount factor, preference and endowment) considered in the paper. This result becomes apparent as we inspect the first-order conditions of an agent’s optimization problem.

#### 3.1 First-Order Conditions

The first-order conditions of System (1) are

\[
\begin{align*}
\beta^t_i u_i'(c_i(s^t)) & = \lambda^i \\
\beta^t_i u_i'(c_i(s^t)) \pi^i(s^t) & = \lambda^i q_0^i(s^t), \quad \forall s^t \in S^t \\
d(s_0) + \sum_{t=1}^T \sum_{s^t} q_0^i(s^t) d(s^t) - c_i(s_0) - \sum_{t=1}^T \sum_{s^t} q_0^i(s^t) c_i(s^t) & = 0
\end{align*}
\]

for \( i = 1, \ldots, I \), where \( \lambda^i \) is agent \( i \)'s Lagrange multiplier for constraint (1b). The intertemporal Euler equation is obtained by dividing (6b) by (6a),

\[
\beta^t_i u_i'(c_i(s^t)) \pi^i(s^t) = q_0^i(s^t),
\]
for all \( s^t \in S^t \) and \( i = 1, \ldots, I \). Selecting agent 1 as a “reference” agent, Equation (7) yields

\[
\frac{\beta^1_t}{\beta^1_1} \frac{u'_i(c(s^t))}{u'_1(c(s^0))} \frac{\pi^i(s^t)}{\pi^1(s^t)} = 1.
\]

(8)

Reformulating (8), we arrive at

\[
c^i(s^t) = u'_i^{-1}\left(\frac{\beta^1_t}{\beta^1_1} \frac{\pi^1(s^t)}{\pi^i(s^t)} \frac{u'_1(c(s^t))}{u'_1(c(s^0))} \right).
\]

(9)

Substituting Equation (9) into the aggregate resource constraint (2),

\[
\sum_{i=1}^I \mu^i(s^t)u'^{-1}_i \left(\frac{\beta^1_t}{\beta^1_1} \frac{\pi^1(s^t)}{\pi^i(s^t)} \frac{u'_1(c(s^t))}{u'_1(c(s^0))} \right) = \sum_{i=1}^I \mu^i(s^t)d(s^t).
\]

(10)

For each \( s^t \in S^t \) and \( t \geq 0 \), given discount rates, \( \{\beta_i\}_{i=1}^I \), beliefs, \( \{\pi^i(s^t)\}_{i=1}^I \), period utilities, \( \{u_i(c)\}_{i=1}^I \), population proportions, \( \{\mu^i(s^t)\}_{i=1}^I \), and initial consumption choices, \( \{c^i(s^0)\}_{i=1}^I \), Equation (10) represents a single nonlinear equation with a single unknown, \( c^i(s^t) \), and is the foundation for understanding trading results for the model under consideration. If \( c^i(s^0) = \bar{c}^i(s^0) \) for \( i = 1, \ldots, I, \) the optimal initial consumption values in competitive equilibrium, the values of \( c^i(s^t) \) and \( \{c^i(s^t)\}_{i=2}^I \) which solve Equations (10) and (9), respectively, will also be the optimal competitive equilibrium values, \( \bar{c}^i(s^t) \) and \( \{\bar{c}^i(s^t)\}_{i=2}^I \), for all \( s^t \in S^t \) and \( t \geq 0 \).

In the general formulation, these optimal choices are history dependent.

### 3.2 Special Cases

For the remainder of this section, let us assume that dividends depend only on the concurrent realization of the aggregate state: \( d(s^t) = d(s_t) \), for all \( s^t \in S^t \) and \( t \geq 0 \). Given the law of motion of \( s^t \), this results in \( d(s_t) \) being Markovian. We will further assume that population proportions are not history dependent: \( \mu^i(s^t) = \mu^i(s_t) \), for all \( i = 1, \ldots, I, s^t \in S^t \) and \( t \geq 0 \). For the results reported in Section 5, we will assume that the population proportions are fixed (\( \mu^i(s^t) = \mu^i \), for all \( i = 1, \ldots, I \), \( s^t \in S^t \) and \( t \geq 0 \), but the following results only require that they are independent of the aggregate state history. Under these restrictions, Equation (10) becomes

\[
\sum_{i=1}^I \mu^i(s_t)u'^{-1}_i \left(\frac{\beta^1_t}{\beta^1_1} \frac{\pi^1(s^t)}{\pi^i(s^t)} \frac{u'_1(c(s^t))}{u'_1(c(s^0))} \right) = \sum_{i=1}^I \mu^i(s_t)d(s_t).
\]

(11)
3.2.1 Homogeneous Beliefs and Time Discount Factors

If we further specialize the model so that agent types have homogeneous beliefs and time discount factors, \( \pi^i(s^t) = \pi(s^t) \) and \( \beta_i = \beta \), for all \( i = \{1, \ldots, I\} \), \( s^t \in S^t \) and \( t \geq 0 \), Equation (11) becomes

\[
\sum_{i=1}^{I} \mu^i(s_t)u_i^{t-1} \left( \frac{u_1'(c^1(s^t))}{u_1'(c^1(s_0))} u_i'(c^i(s_0)) \right) = \sum_{i=1}^{I} \mu^i(s_t)d(s_t). \tag{12}
\]

For a given set of initial consumption choices \( \{c^i(s_0)\}_{i=1}^{I} \), the value of \( c^1(s^t) \) satisfying Equation (12) and the associated values of \( \{c^i(s^t)\} \) obtained from (9) will only depend on \( s_t \) and not on \( s^{t-1} \). This is true even if agent types have heterogeneous period utility functions or endowments, so long as those variables do not depend on state histories themselves.

The implication of Equation (12) is that the associated competitive equilibrium consumption choices will be Markovian and time homogeneous; if \( s_t = s \in S \), then \( \bar{c}^i(s^t) = \bar{c}^i(s) \), independent of time and for all \( i = 1, \ldots, I \). This is the key result of Judd et al. (2003), and, as noted in their paper, hinges on the Markovian structure of the economy (notably, dividends). For general complete asset markets, where agents achieve their desired state-contingent consumption levels through a mix of asset purchases, this carries important trade implications, as argued by Judd et al. (2003). For example, in the case where agents have access to \( S \) infinitely-lived assets, agent types will trade once at \( t = 0 \) and never again. Allowing for finite-lived assets generates trade at periods \( t > 0 \), but only so that agents can hold constant portfolios.

3.2.2 Heterogeneous Time Discount Factors

Maintaining homogeneity of beliefs, but allowing agents to have varying time discount factors yields

\[
\sum_{i=1}^{I} \mu^i(s_t)u_i^{t-1} \left( \left( \frac{\beta_1}{\beta_i} \right)^t \frac{u_1'(c^1(s^t))}{u_1'(c^1(s_0))} u_i'(c^i(s_0)) \right) = \sum_{i=1}^{I} \mu^i(s_t)d(s_t). \tag{13}
\]

This version of the model has properties similar to that of the previous subsection, except that the optimal consumption choices now depend on the date \( t \), through the ratio of time discount factors, \( \left( \frac{\beta_1}{\beta_i} \right)^t \). In this sense, the consumption policy rules will depend on the
history of states in a deterministic way: they will not depend on a particular history \( s^t \)

itself, but on the number of periods which have passed. It is easily shown that in such an
economy, the most patient individual (with the highest time discount factor) will trade so
that he consumes the full dividend in the distant future while all other agent types consume
nothing. In effect, the most patient agent survives. The speed with which this happens is
increasing in the dispersion of time discount factors, but even for very large discrepancies,
the trading volume between periods is quite small and follows deterministic patterns that
lead to survival of the most patient agent type.

3.2.3 Heterogeneous Beliefs

Returning to the case of heterogeneous beliefs, it is apparent from Equation (11) that optimal
consumption choices \( \{c^i(s^t)\}_{i=1}^I \) will not be time homogeneous, even when dividends and
population proportions are not history dependent. The reason is that, even when \( s^t \) is
Markovian, the agent specific transition probabilities introduce a history dependence that
cannot be eliminated; since each agent type maintains distinct beliefs through time, the path
of the economy is no longer trivial in their relative probability assessments.

The important result, then, is that asset trading will not be trivial in equilibrium, even
in the presence of complete markets. In the case of \( S \) infinitely-lived financial assets, the
no-trade equilibrium will not be a general result. Similarly, with assets of varying finite
maturities, agent types will not necessarily hold constant portfolios. This particular result
is developed in Section 5.

3.3 An Equivalent Problem

It is possible to recast the general problem with history dependence as one of history-
dependent time discount factors:

\[
\sum_{i=1}^I \mu^i(s_t)u_t^{-1}\left(\frac{\tilde{\beta}^i_t(s^t)u_t'(c^1(s^t))}{\tilde{\beta}^i_t(s^t)u_t'(c^1(s^0))}u'_i(c^i(s_0))\right) = \sum_{i=1}^I \mu^i(s_t)d(s_t),
\]

where \( \tilde{\beta}^i_t(s^t) = \beta^i_t \frac{\bar{\pi}(s^t)}{\bar{\pi}(s^0)} \). In this case, we are viewing the heterogeneous beliefs model as
one with distorted homogeneous beliefs, \( \bar{\pi}(s^t) \), and state- and history-dependent time dis-
count factors. The non-deterministic history-dependent time discount factors, \( \tilde{\beta}^i_t(s^t) \), are
able to generate the same trading patterns that we observe in an alternate economy with heterogeneous beliefs, $\pi^i(s^t)$, and heterogeneous, but fixed, time discount factors, $\beta_i$.

4 Solution Method

For the remainder of the paper, we will focus on a finite-horizon version of the model in Section 2. This has no implication for the theoretical arguments developed in Section 3, but facilitates computation of the competitive equilibrium when optimal consumption choices are not homogeneous. Consider the following algorithm.

1: Fix some $\tau > 0$, which will determine convergence and set $\varepsilon = 1$.
2: Guess initial values for $c^i(s^0)$, $i = 2, \ldots, N$.
3: for $t = 1, \ldots, T$
4: for $s^t \in S^t$
5: Determine $c^1(s^t)$ such that

$$\sum_{i=1}^I \mu^i(s^t) u_i^{-1} \left( \frac{\beta_i^t \pi^1(s^t) u_1'(c^1(s^t))}{\beta_1^t \pi^1(s^t) u_1'(c^1(s^0))} u_i'(c^i(s^0)) \right) \right) = \sum_{i=1}^I \mu^i(s^t)d(s^t). \quad (15)$$

6: for $i = 2, \ldots, I$
7: Compute

$$q^0(s^t) = \beta_1^t u_1'(c^1(s^t)) \pi^1(s^t) \quad (16)$$

$$c^i(s^t) = u_i^{-1} \left( \frac{\beta_i^t \pi^1(s^t) u_1'(c^1(s^t))}{\beta_1^t \pi^1(s^t) u_1'(c^1(s^0))} u_i'(c^i(s^0)) \right). \quad (17)$$

8: end for
9: end for
10: end for
11: for $i = 2, \ldots, N$
12: Compute

$$\varepsilon^i = d(s^0) + \sum_{t=1}^T \sum_{s^t} q^0(s^t)d(s^t) - c^i(s^0) - \sum_{t=1}^T \sum_{s^t} q^0(s^t)c^i(s^t). \quad (18)$$

13: end for
14: Set $\varepsilon = \sum_{i=2}^{N} |\varepsilon^i|$.  
15: if $\varepsilon > \tau$ then  
16: Use Broyden’s method to choose new values of $c^i(s_0)$ and return to Step 3.  
17: else  
18: Stop.  
19: end if  

4.1 GPU Computing  

The algorithm above is well suited for parallel computing. For a $T + 1$ period economy, there are a total of $\frac{s^{T+1}-1}{s-1}$ states: $s^t$ at each time period $t = 0, 1, \ldots, T$. Figure 1 depicts an example with $S = 2$ and $T = 3$. Given a candidate $\{c^i(s_0)\}_{i=2}^{I}$, Equations (15) – (17) determine a solution for $q^0(s^t)$ and $\{c^i(s^t)\}_{i=1}^{I}$ at each of the $\frac{s^{T+1}-1}{s-1} - 1$ states after $t = 0$, independent of the prices and consumption choices at other dates and states in the economy. This independence between states allows the computation to be divided into distinct pieces, each of which can be performed by a separate processing unit. In theory, with enough processors, it would be possible to assign the nonlinear equation problem (Equation (15)) of each state to one processor. In practice, with large $S$ or large $T$, a subset of state-tree nodes would be assigned to each processor for computation.

The inherent parallelism in the above algorithm is ideal for graphics processing unit (GPU) computing. As documented by Aldrich et al. (2011), modern graphics cards typically contain hundreds, sometimes thousands, of individual processing cores. These cores are optimized for performing simple sets of instructions in parallel, and are not ideally suited for algorithms which require heavy memory use or many branches. For the algorithm considered above, the actual computational problem that needs to be solved at each state is relatively simple: it amounts to solving a single equation for a single variable, using Newton’s method with a known and analytically simple derivative. With hundreds or thousands of individual cores to solve these nonlinear equations, GPU computing can deliver enormous speed gains over traditional CPU computing for values of $S$ and $T$ that result in a large state space $S$. 

11
5 Empirical Application

The objective of this section will be to parameterize the model so that it simultaneously captures broad features of quarterly aggregate consumption growth as well as quarterly equity market volume and price changes. This section provides a brief overview of the aggregate data that is relevant to the model, discusses estimation of the endowment process parameters, reports implications of the model and concludes with a discussion of the economic motives underlying the results.
5.1 Data

To both estimate the model and to provide a reference for model output, we consider data from three different sources:

1. Real, per capita, quarterly aggregate consumption (Nondurables and Services) from 1947.1 to 2010.4 (NIPA Table 7.1).

2. Monthly volume and shares outstanding for all assets traded on the NYSE from March 1947 to December 2010 (CRSP Monthly Stock Files).


The data period was chosen to begin in 1947, because this is the oldest quarterly consumption data available from the BEA. Individual volumes and shares outstanding were summed across all assets and over relevant months to obtain a measure of aggregate quarterly turnover, where

\[
\text{Turnover} = \frac{\text{Volume}}{\text{Shares Outstanding}}. \tag{19}\]

Technically speaking, volume refers to the number of shares traded within a given period, not the proportion of shares traded (turnover). Since the number of shares in the model is normalized to unity, these quantities are identical and the terms will be used interchangeably for the remainder of the paper.

Figure 2 displays time series plots of quarterly aggregate consumption growth, quarterly NYSE turnover and quarterly absolute price changes for the S&P 500 (quarterly changes were obtained by differencing index levels at three month intervals). It is readily apparent from the figure that turnover and price changes have been trending upward since the early 1980s, with episodes of large price changes occurring in the late 1990s and early 2000s, and a burst of turnover surrounding the recent financial crisis. Figure 3 highlights a well-documented feature of price-volume data: that absolute changes in price are positively correlated with turnover. This empirical relationship has been observed at daily, weekly, monthly, quarterly and annual frequencies, as documented by Karpoff (1987). Figure 3 distinguishes data
Figure 2: Time series plot of real, per capita, aggregate consumption growth, aggregate NYSE turnover and S&P 500 absolute price changes. All series are quarterly, 1947.2 to 2010.4.

pre- and post-2000, the former represented with empty circles and the latter with circles superimposed with crosses. Interestingly, while a strong positive relationship exists before 2000, an even stronger relationship is apparent in the data after 2000.

Finally, Figure 4 depicts a kernel density estimate of the NYSE turnover series. The mean and median of quarterly NYSE turnover are 0.1542 and 0.08605, respectively. Further, more than 75% of the density lies below the value 0.2 and more than 90% lies below 0.4. These values provide a reference target for a model of quarterly market volume.

5.2 Estimation

We now consider a specialization of the model in Section 2 with $S = 2$ and $I = 2$. We will assume that proportions of agent types are fixed through time, $\mu^i(s^t) = \mu^i$, for $i = 1, 2$, and
that agents have constant relative risk aversion utility,

$$u^i(c) = \frac{c^{1-\gamma^i}}{1-\gamma^i}. \tag{20}$$

Agents receive aggregate consumption, $C(s^t)$, as their endowment in state $s^t$, where the two aggregate states of nature each period, $s_t$, represent high consumption growth and low consumption growth ($s_t \in \{l,h\}, \forall t$). Following Hamilton (1989), we specify a hidden Markov model for the logarithm of aggregate consumption growth,

$$\Delta c_t = \alpha(s_t) + \varepsilon_t, \tag{21}$$

where $\Delta c_t = \log(C_t) - \log(C_{t-1})$, $\alpha(h) > \alpha(l)$, and where $s_t$ obeys a Markov transition law, $\pi(s_t|s_{t-1})$, as described in Section 2. Table 1 reports maximum likelihood estimates and standard errors for the hidden Markov parameters. The parameter estimates suggest that high and low growth states have roughly the same magnitude within the series, but that

Figure 3: Scatter plot of absolute price changes for the S&P 500 and aggregate NYSE volume. Both series are quarterly, ranging from 1947.2 to 2010.4. Empty circles represent data values before 2000 and the circles superimposed with crosses represent post-2000 data.
Figure 4: Density estimate of quarterly aggregate NYSE volume from 1947.2 to 2010.4. The density was computed with a Gaussian kernel and a bandwidth of 0.03244.

high growth is much more persistent than low growth, which has a roughly equal probability of transitioning to a subsequent high or low growth state.

Similar to Cecchetti et al. (2000) and Cogley and Sargent (2009), we will eliminate the disturbance in Equation (21) and force aggregate consumption growth in the model to be equal to the estimated means. That is,

$$C(s^{t+1}) = g(s^{t+1})C(s^t),$$

where $g(s^t) = \exp(\alpha(s_t))$. This specification causes the aggregate endowment sequence to be history dependent. In the notation of the previous sections, $C(s^t) = d(s^t)$ is dependent on the full history $s^t$ and not just the concurrent state of nature $s_t$. The general implication of this state dependence is that both preference (utility) heterogeneity and endowment heterogeneity can result in consumption rules that are not time homogeneous, and which lead agents to hold non-constant portfolios (i.e. to trade at $t > 0$). However, several caveats are
important to note. First, when agents have power utility and identical coefficients of relative risk aversion (as well as identical discount factors and beliefs), their optimal consumption choices are constant fractions of the aggregate endowment (as shown below in Equations (29) and (30)), leading to no asset exchange after $t = 0$. Second, with power utility and heterogeneous coefficients of relative risk aversion, agents engage in trade at $t > 0$, but the total volume of trade is extremely small (results are not reported here for space considerations). Hence, it is not problematic to investigate the trade implications of heterogeneous discount factors and beliefs in this setting with state dependent dividends.

5.3 Quantifying Belief Divergence

In the model, the majority of agents will maintain beliefs that are consistent with the estimated probabilities in Table 1, while a minority of the population will deviate. In particular, the following section will consider cases where the minority believes $\pi(l|l)$ is one, two and three standard errors below its maximum likelihood estimate; hence, they are relatively optimistic. Before considering the implications of such heterogeneity, let us first quantify the magnitude of the disturbances and determine if the proposed belief divergence is plausible.

Table 2 reports annualized moments of the simulated log consumption growth process using the proposed values of $\pi(l|l)$ as well as unconditional quarterly probabilities of recession and expansion. As is expected, optimism (lower values of $\pi(l|l)$) is associated with higher mean consumption growth and lower volatility as well as lower unconditional probabilities of recession and higher unconditional probabilities of expansion. As an example, an agent whose views deviate from the baseline (MLE) case by two standard deviations, believes that consumption will grow an average of 0.16% more annually and will have 0.1124% lower annual
volatility. In addition, such an agent believes a recession (expansion) to be 3.5% less (more) likely each quarter. These values provide a simple and intuitive measure of the magnitude of belief disturbance and are important to understand when interpreting the model results.

More formally, we can quantify belief divergence with detection error probabilities, a notion developed by Anderson et al. (2000) and Hansen et al. (2002), with roots in Chernoff (1952). In particular, denote the baseline probability model as $B$ and the distorted (optimistic) model as $D$. Given a $T$-period observed sequence of consumption values, $C^T \equiv C(s^T)$, we can evaluate the likelihood of the data under each of the probability models: $L^B(C^T)$ and $L^D(C^T)$. Given that agents believe probability model $i \in \{B, D\}$ is correct when the likelihood is greater, we can determine the probability of making a mistake under each of the models:

$$p^B \equiv P(\text{mistake}|B) = P(L^D(C^T) > L^B(C^T)|B)$$

$$p^D \equiv P(\text{mistake}|D) = P(L^B(C^T) > L^D(C^T)|D).$$

Table 2: Annualized moments of simulated log consumption growth for various values of $\pi(l|l)$ and unconditional (quarterly) probabilities of recession and expansion. Simulations consist of 1 million observations.

|       | Data           | $\pi(l|l)$          |
|-------|----------------|---------------------|
|       |                | 0.5304 (MLE) 0.4091 0.2878 0.1665 |
| Mean  | 0.01949        | 0.01988 0.02082 0.02148 0.002190 |
| SD    | 0.01012        | 0.007072 0.006449 0.005948 0.005600 |
| $\pi(l)$ | NA          | 0.1114 0.09064 0.07638 0.06600 |
| $\pi(h)$ | NA          | 0.8886 0.9093 0.9236 0.9340 |
made; the former is an approximate measure while the latter is exact. The detection error probability for horizon $T$ is then determined by

$$p = \frac{1}{2} (p^B + p^D), \quad (25)$$

which is interpreted as the posterior probability of a mistake, given an equally weighted prior distribution on the two models. Clearly, the maximum value of the detection error probability is

$$\lim_{B \to D} p = \frac{1}{2} \left( P(L^D(C^T) > L^B(C^T)|B) + P(L^D(C^T) < L^B(C^T)|B) \right)$$

$$= \frac{1}{2} (p^B + (1 - p^B))$$

$$= \frac{1}{2}. \quad (26)$$

That is, with very little data or as the models approach each other and become indistinguishable, we observe a maximum detection error probability of 0.5. Alternatively, with an infinite amount of data or when the models are far from each other, the probability of a mistake approaches zero.

Figure 5 plots detection error probabilities when $T = 12$ quarters, for the baseline MLE model above and distorted models with probabilities $\pi(l|l)$ that decline from the MLE toward zero. The dotted lines mark the detection error probabilities of each of the distorted models considered in the next section ($\pi(l|l)$ one, two and three standard deviations below the MLE). The associated error probabilities corroborate the intuition provided by the consumption simulations in Table 2: the high probabilities suggest that under each of the distorted models, agents have very little ability to distinguish the true model with only 12 quarters of observed data. That is, the probability of a mistake is quite high. This is important, since the ability to generate trading volume and appropriate price movements would be of little interest in a framework where agents can easily detect that the distorted model is false.

5.4 Model Implications

To understand the impact of heterogeneous beliefs on asset exchange over a finite horizon, we will consider an economy with $T + 1 = 12$ periods (or three years). Let us assume that
\[ 
\pi(l|l) = 0.4091 \\
\pi(l|l) = 0.2878 \\
\pi(l|l) = 0.1665 
\]

**Figure 5:** Detection error probabilities for \( T = 12 \) quarters under MLE and increasingly optimistic models (\( \pi(l|l) \) declining to zero).

\( \mu^1 = 0.8 \) and \( \mu^2 = 0.2 \), where, as described above, type 1 agents (the majority) possess beliefs that are parameterized by the hidden Markov maximum likelihood estimates, and type 2 agents believe \( \pi^2(l|l) < \pi^1(l|l) \). In addition, assume \( \gamma^1 = \gamma^2 = 2 \) and \( \beta = 0.995 \). The latter value corresponds to an annualized risk-free rate of 2.03%.

Table 3 reports trading volume summary statistics for each of the long-lived assets, with the values of \( \pi(l|l) \) considered above. In each case, the model was solved using the methods of Section 4 and means and standard deviations were computed using the MLE transition probabilities in Table 1 to compute \( \pi(s^t) \); that is, moments were computed under the assumption that type 1 agents hold correct beliefs. Trade at \( t = 0 \) is excluded from the summary statistics since it is highly dependent on the initial endowment and can be
manipulated to fabricate desirable volume statistics.

When beliefs are homogeneous, agents do not exchange assets (beyond $t = 0$) and trading volume is identically zero. However, Table 3 indicates that, depending on the degree of belief heterogeneity, trading volume can reach values as high as 20% of outstanding shares, and can average up to 6% of the market. The reported statistics also indicate that volume varies substantially across the 12 time periods, as encapsulated by standard deviations of 0.6 - 5.1%, with the variability positively related to degree of heterogeneity. Even a moderate degree of heterogeneity ($\pi^2(l|l) = 0.2878$) can generate a maximum value for volume that exceeds 12%, which is roughly 80% of the average quarterly NYSE turnover reported above (15.42%). Although these values fall a little under those found in the data, they are qualitatively sensible and are a substantial achievement in a complete-markets general-equilibrium framework that captures the movement of aggregate consumption.

Most compelling is that the model is not only able to generate sizable volume, but is able to generate the right relationship between volume and absolute price changes. Figure 6 depicts a scatter plot of expected volume against expected absolute prices changes, for the intermediate case of $\pi^2(l|l) = 0.2878$, using the baseline (MLE) probabilities for each state. The broad relationship is a success: as documented in the empirical literature, volume exhibits positive correlation with absolute price changes. The next section will explain why heterogeneous beliefs are instrumental for this relationship, and hence why they are an important component of asset exchange in financial markets.

Table 4 reports similar trading volume statistics for the case of heterogeneous discount factors: $\beta_1 = 0.995$ and $\beta_2 \in \{0.99, 0.985, 0.98\}$. The values of $\beta_2$ roughly correspond to annual risk-free rates of 4.1%, 6.2% and 8.4%; in other words, very optimistic views
(relative to historical data) about risk-free returns. Even with such optimistic views, Table 4 shows that discount factor heterogeneity delivers very little asset exchange, in all cases never exceeding 0.3% of the market, with standard deviations that are uniformly less than 0.1%. Figure 7 depicts a scatter plot of expected volume against expected absolute prices changes for the case when $\beta_2 = 0.985$. Unlike the heterogeneous beliefs case, the model does not generate a positive relationship between the two variables.

### 5.5 Speculation and Trading Volume

We now consider agents’ motivation to engage in trade and the relationship of trade volume and price changes to heterogeneous beliefs. We begin with some useful mathematical simplifications.
<table>
<thead>
<tr>
<th>Asset</th>
<th>( \beta_2 = 0.99 )</th>
<th>( \beta_2 = 0.985 )</th>
<th>( \beta_2 = 0.98 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Min</td>
</tr>
<tr>
<td>Asset 1</td>
<td>0.00070</td>
<td>0.00024</td>
<td>6.659e-06</td>
</tr>
<tr>
<td>Asset 2</td>
<td>0.00070</td>
<td>0.00023</td>
<td>6.659e-06</td>
</tr>
</tbody>
</table>

**Table 4:** Trading volume for a 12-period complete-markets endowment economy, excluding trade at \( t = 0 \). Separate statistics are reported for each case of \( \beta_2 \).

When agents have constant relative risk aversion utility, Equation (9) simplifies to

\[
c^i(s^t) = \xi^i(s^t)c^1(s^t)^{\gamma^i/\gamma^1}
\]  
(27)

where

\[
\xi^i(s^t) = \left( \frac{\beta_2^i}{\beta_1^i} \pi_i(s^t) \right)^{1/\gamma^i} \frac{c^i(s_0)}{c^1(s_0)^{\gamma^1/\gamma^i}}.
\]  
(28)

Furthermore, when agents have identical beliefs, discount factors and coefficients of relative risk aversion,

\[
\xi(s^t) = c^i(s_0)/c^1(s_0),
\]

which is history independent (it only depends on \( s_0 \)), and from Equations (10) and (9),

\[
c^1(s^t) = \tau(s_0)d(s^t)
\]  
(29)

and

\[
c^i(s^t) = \xi^i(s_0)\tau(s_0)d(s^t)
\]  
(30)

where

\[
\tau(s_0) = \left( \sum_{i=1}^{I} \mu_i \xi^i(s_0) \right)^{-1}.
\]  
(31)

The implication is that when agents are homogeneous (aside from endowments) and have power utility, their optimal consumption choices are a constant fraction of the aggregate dividend.
Simsek (2011) introduces a notion of speculation that is encapsulated by the volatility of agents’ net worth: he shows that the introduction of new assets in an incomplete market with heterogeneous beliefs creates a motive for agents to both hedge and speculate on aggregate risk. The degree of hedging and speculation is captured by the contribution of each to the volatility of agents’ net worths. Intuitively, when agents fully hedge themselves, the only volatility in their net worths should be attributed to the volatility of aggregate risk.

In our complete-markets economy, when agents are homogeneous, they fully hedge themselves by choosing time-homogeneous consumption plans that lead to constant net worths, as a fraction of the aggregate endowment (which controls for aggregate variability). That is, according to Equation (30)

$$\omega^i(s^T) = \sum_{t=0}^{T} \beta_t^i \xi^i(s^t) d(s^t) = \sum_{t=0}^{T} \beta_t^i \xi(s_0) \tau(s_0),$$

(32)
is constant for all $s^T \in S^T$, given $s_0 \in S$. However, the introduction of heterogeneous beliefs or discount factors causes the fractions $\xi^i(s^t)$ in Equation (28) to become history dependent. This causes agents’ net worths, expressed as a fraction of the aggregate dividend, to be history dependent, and to contain variability other than that of the aggregate state. Tables 5 and 6 summarize the volatilities of net worths for the various values of $\pi^2(l|l)$ and $\beta_2$ considered above. Interestingly, heterogeneous discount factors introduce very little variability in

|                | $\pi^2(l|l) = 0.4091$ | $\pi^2(l|l) = 0.2878$ | $\pi^2(l|l) = 0.1665$ |
|----------------|-----------------------|-----------------------|-----------------------|
| Type 1 Type 2  | 0.4081 1.632          | 0.8081 3.232          | 1.184 4.737          |

Table 5: Volatility of net worths across all possible state histories in a 12-period complete-markets endowment economy. Separate statistics are reported for each case of $\pi^2(l|l)$.

<table>
<thead>
<tr>
<th></th>
<th>$\beta_2 = 0.99$</th>
<th>$\beta_2 = 0.985$</th>
<th>$\beta_2 = 0.98$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1 Type 2</td>
<td>8.299e-06 9.435e-06</td>
<td>1.050e-05 7.111e-06</td>
<td>8.112e-06 7.945e-06</td>
</tr>
</tbody>
</table>

Table 6: Volatility of net worths across all possible state histories in a 12-period complete-markets endowment economy. Separate statistics are reported for each case of $\beta_2$.

net worth. Heterogeneous beliefs, on the contrary, introduce substantial variability. The reason for this is that diverse beliefs motivate agents to speculate against each other: type 1 agents choose higher values of consumption on state paths that they view as relatively more likely than their counterparts, and vice versa. This intuition is captured in Figure 8, which plots net worths against log probability ratios of entire histories, $s^t$. The size of the circles represents the magnitude of agents’ absolute probability assessment, independent of the opposing agents’ beliefs. The figure clearly depicts that agents increase their net worth on paths of the economy that correspond to high relative probability, as measured by the counterpart’s beliefs. In addition, absolute probability assessment appears to have no relationship with net worth.
Figure 8: Scatter plot of log likelihood ratios and net worths (as a fraction of the aggregate dividend) for a 12-period endowment economy with heterogeneous beliefs. Values are computed for each state. Circle size represents the magnitude of absolute probability assessment.

We now begin to understand the relationship between trading volume and beliefs: heterogeneous beliefs create a motive for agents to speculate against each other on certain paths of the aggregate state, which induces them to hold non-constant portfolios of long-lived assets and to engage in trade at periods $t > 0$. The same is not true of other forms of heterogeneity. In addition, the magnitude of speculation and trading volume is closely related to the relative probability assessment; it is clear from Equation (28) that $\xi^i(s')$ is positively related to the likelihood ratio $\pi^i(s')/\pi^1(s')$ and inversely related to the coefficient of relative risk aversion.

The notion of speculation is also captured by observing the behavior of prices when an
agent with divergent beliefs is introduced into an otherwise homogeneous economy. For example, the first panel of Figure 9 depicts a scatter plot of log price ratios against log probability ratios, state by state, under the two scenarios where the economy is entirely populated by type 1 agents, and where the economy is populated by both types (with the proportions indicated above). The quantities of the former scenario (type 1 homogeneity) are the divisors. These ratios can be interpreted as the percentage changes in probabilities and prices when a small portion ($\mu^2 = 0.2$) of type 1 agents in a homogeneous economy suddenly transition to become type 2 agents. Similarly, the second panel of the figure depicts the case when a large portion ($\mu^1 = 0.8$) of type 2 agents in a homogeneous economy suddenly transition to become type 1 agents. It is apparent from Figure 9 that prices increase in states

![Figure 9](image-url)

**Figure 9:** Scatter plot of log likelihood ratios and log price ratios. Divisors in the ratios correspond to quantities in a homogeneous economy of the specified base type, and numerators correspond to the equivalent quantities in the heterogeneous economy. In each case, $T + 1 = 12$. 

27
to which divergent agents assign a high relative probability. This makes intuitive sense, as those agents seek to trade additional shares of the state-contingent asset in order to increase their consumption along those paths. The curvature apparent in the first panel of the figure is related to the relative magnitudes of the agents and the scale of the price changes. The same curvature exists in the second panel, but is not perceptible due to the magnitude of price changes, which in turn is directly related to the large quantity of divergent agents introduced in the baseline scenario.

This latter notion of speculation is related to Harrison and Kreps (1978). In their work, Harrison and Kreps (1978) show that agents with heterogeneous beliefs are willing to purchase assets at prices higher than their intrinsic worth, simply because they anticipate selling them for a capital gain to other agents at a later period. They refer to this as a speculative motive, since it is driven by the bet of selling for a capital gain, and show that equilibrium prices are higher in all states, for all assets. Their development, however, hinges on a combination of risk neutrality and no short sales constraints, which causes only one agent type to hold the single asset at any point in time. The model of this paper shows similar results for a more general framework with complete markets and risk-averse agents. In particular, agents push prices above their homogeneous economy counterparts on state paths that they view as relatively more likely. In this sense, pessimistic agents are willing to purchase assets at higher prices along paths that optimists view as more likely, because they have a greater chance of selling those assets to optimists (who value them more highly) along those paths.

In summary, heterogeneous beliefs induces a single motive for agents to alter their consumption on paths that they view as relatively more likely. However, the consequence of this single motive is that both asset exchange (volume) and prices increase on those paths. This results in a positive relationship between absolute prices changes and trading volume, as observed in empirical data. Thus, not only can belief heterogeneity account for empirically reasonable levels of market volume within a general-equilibrium framework, it also captures proper behavior in relation to other market variables. This makes belief heterogeneity a very appealing candidate for explaining both asset exchange as well as other asset market phenomena, especially within a macroeconomic framework.
6 Conclusion

This paper has shown that belief heterogeneity is inherently different from other forms of standard endowment, preference and discount-rate heterogeneity with respect to trade implications for a general-equilibrium complete-markets endowment economy. Judd et al. (2003) showed that in the presence of these other forms of heterogeneity, an important no-trade result holds: when markets are complete and agents have access to a spanning set of infinitely-live assets, they engage in trade at $t = 0$ and never again; with a spanning set that includes finite-maturity assets, agents hold constant portfolios which allow them to maintain a time-homogeneous consumption path.

Heterogeneous beliefs, however, introduce a history dependence that leads agents to desire consumption paths that are not time homogeneous. In essence, agents bet on realizations of state histories that they view as relatively more likely, when compared to their counterparts. This speculative motive is reflected in the variability of agents’ net worths, which are constant across history realizations when agents have homogeneous views and completely hedge themselves. The upshot is that agents’ desire to increase consumption on certain paths while decreasing it on others (i.e. a lack of time homogeneity), fuels nontrivial trading at dates beyond $t = 0$, even when markets are complete. This motive also puts pressure on asset prices and generates a positive relationship between absolute price changes and trading volume.

This paper also demonstrates the trade implications of heterogeneous beliefs in a finite-period application that is congruent with aggregate historical data. In the application, consumption growth states and transition probabilities are calibrated to match an estimated hidden Markov model for post-war quarterly consumption growth. A small proportion of agents maintain optimistic views (of varying degrees) regarding transitions from a low consumption growth state to a high consumption growth state. Such a specification leads to sizeable trading volume (as high as 20% of shares outstanding) which has a positive relation to absolute price changes, as documented in the empirical literature. Other forms of heterogeneity, such as discount factor heterogeneity, can only generate an extremely low volume of trade (almost negligible) for very optimistic parameterizations, and also cannot obtain
proper qualitative relationships with prices.

Finally, this paper suggests an algorithm for solving the finite-period model with heterogeneous beliefs and demonstrates its suitability for large-scale parallel computing. The optimal allocations in competitive equilibrium are chosen at each state, independent of all but the initial state of the economy. The resulting mathematical structure, then, is one that can leverage many processing units to compute optimal consumption plans for many periods and many states. In particular, advances in GPU computing, as suggested by Aldrich et al. (2011), can be utilized to compute the model with more agents, states and periods than would be feasible in a traditional CPU framework.

References


