Industrial Asset Pricing with Endogenous Business Cycles*

Marcus M. Opp Christine Parlour Johan Walden

March 1, 2012

Abstract

We develop a dynamic general equilibrium model with a continuum of different industries, each comprising a finite number of strategic price-setting firms. In our framework, the market price of risk is endogenous and determined by the consumption of a representative agent. General equilibrium in the model is shown to exist under general conditions. Strategic interaction between firms amplifies business cycles, and the equilibrium outcome can be very sensitive to small changes in long-term growth rates whereas temporary changes have only marginal impact. A firm's expected returns are affected by the industrial environment in which it operates, in line with what has been observed in the empirical literature. Overall, our model suggests that strategic competition may drive business cycles and that industry characteristics should be informative about the expected returns of individual firms.

*Preliminary version. Comments are welcome.** We thank Kyle Bagwell, Nicolae Garleanu, Zhiguo He, Anders Karlsson, Debbie Lucas, Hanno Lustig, and Robert Staiger for very helpful comments and discussions. We also appreciate feedback from seminar participants at Aarhus University, Goethe University, University of Illinois, Urbana Champaign, University of Southern California, and the 2012 Jackson Hole finance conference.
All authors are at the Haas School of Business, UC Berkeley.

1 Introduction

In 1973 the OPEC cartel raised oil prices, precipitating a global recession. This event was unique but not unprecedented: Several of the largest shocks in financial markets as well as in the real economy in the last century originated from the strategic behavior of firms and other entities. At the industry level, strategic interactions, e.g., through collusion, price wars and new entry, fundamentally determines the profitability and value creation of firms, and shocks are the norm rather than the exception. The general equilibrium asset pricing literature does not speak to such strategic interactions. Typically, in this literature, shocks to the economy are "technological," driven by exogenous processes. However, history suggests that technological shocks, although important, do not capture the whole story. Understanding the general equilibrium effects of such strategic behavior should therefore be an important goal for asset pricing and macro economics more broadly.

We analyze a discrete time, infinite horizon general equilibrium model with many different industries, in which firms within each industry interact strategically in a product market. Specifically, we embed a dynamic oligopoly game in each industry into a general equilibrium model with a continuum of production technologies similar to the seminal paper by Rotemberg and Woodford (1992). Our key departure from their paper is the introduction of sector heterogeneity (in terms of productivity and the number of players) which drives all of our novel implications, in particular endogenous business cycles driven by asynchronous behavior of industries. Within each industry, we model price competition as the subgame perfect outcome in which firms choose their markups. In each period, each firm weighs the value of high short-term profits that can be obtained by aggressive pricing, against the long-term profits that are obtained when firms cooperate. All profits are valued in a standard consumption based asset pricing framework by a representative agent who consumes the goods and values future (risky) cash flows. While each industry takes the macro dynamics, in particular aggregate consumption, as given, industries jointly affect these macro dynamics in a non-trivial way, as long as the different sectors' equilibrium strategies are not perfectly aligned.

Most of the intuition in our paper comes from the interplay between asset pricing and industrial organization. This is because aggregate consumption is endogenous and depends on the output in each industry and also, through the pricing kernel, affects the ability of firms in each industry to sustain collusive outcomes. A small productivity shock in one industry can have an economy-wide effect because firms' profits, values and returns therefore vary with the changing competitive environment. The effects of strategic interaction is important in the sense that small technological shocks can have a drastic impact on the equilibrium outcome and on asset prices, through the mechanism of strategic competition.

Our paper makes three types of contributions. First, given that the setting with multiple industries and strategically interacting firms is fairly complex, we spend substantial effort on developing a rigorous understanding of the model, deriving general equilibrium existence results and exploring the properties of the model. Our main result in this part of the paper is Proposition 5, which shows the existence of equilibrium under minimal assumptions.

Second, we provide several results that show the effect of strategic interaction on the economy's equilibrium. Specifically, strategic interaction can amplify, and even generate, business cycles endogenously. The welfare costs of such interaction are the highest in economies in which competitiveness varies a lot across industries, whereas they are low when the variation is low, even in the case when all industries are monopolistic. Further, business cycle fluctuations are most severe in economies in which there is considerable dispersion across industries' in the ability to sustain high markups when aggregate productivity is low. The equilibrium outcome can be very sensitive to small changes in long term growth-rates whereas temporary changes, on the other hand, have at most a marginal impact. Finally, the economy may have multiple qualitatively very different equilibria. There is therefore a role for policy in our framework.

Third, we study asset pricing implications of the model, relating industry- and economy wide-characteristics with expected returns in the stock market. For the benchmark case with monopolistic firms, the results are quite straightforward. Under some technical conditions, firms in industries with pro-cyclical product demand have higher expected returns than firms in countercyclical industries. This is in line with the intuition that pro-cyclical firms generate value in good states of the world, in which marginal utility is low, and that such firms are therefore discounted at a higher discount rate. When there is strategic competition, however, this intuition breaks down, and non-monotonic relationships between product demand and expected returns may occur. The general implication is that industry characteristics (product demand, industry concentration and markups) should be informative about a firm's expected returns in the stock market.

Our paper is related to the Industrial Organization literature on strategic competition over the business cycle (see, e.g., Bagwell and Staiger, 1997). Our model takes this literature as a starting point, but extends the approach to allow for multiple industries and endogenous pricing of risk in an economy with risk averse agents. Our paper is also related to the extensive business cycle literature (Kydland and Prescott, 1982; Long and Possner, 1983; Gabaix, 2011; Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi, 2011). Specifically, our model highlights how strategic interaction between firms (as an alternative to technological shocks) can generate endogenous business cycle fluctuations. It may also lead to inefficient outcomes, suggesting a role for policy makers. Our paper is further related to a small, but growing finance literature that explores the relationship between firms' strategic environment in the product market and their expected returns in the stock market (Hou and Robinson, 2006; Hoberg and Phillips, 2010; Lyandres and Watanabe, 2010), which find that industry concentration is informative about a firm's expected returns. Hou and Robinson (2006) document that firms in highly concentrated industries earn lower returns. Hoberg and Phillips (2010) find stronger "boom-bust" cycles in competitive industries, in that historically high valuations in competitive industries are stronger predictors of negative abnormal returns than in more concentrated industries. Lyandres and Watanabe (2010) suggest that cross sectional variations in industry concentration can explain the size and value premiums. These effects are all consistent with our model. Finally, our paper is related to the literature that use additional balance and income sheet items, beyond proxies for free cash flows, to explain expected returns (see, e.g., the discussion in Novy-Marx (2010)). Our model provides a motivation for why such additional items matter.

To get the many different pieces of the model to fit together in a consistent and tractable framework, we have been forced to think carefully about modeling choices. To have a strategic trade-off among firms between competition and cooperation, multiple periods are needed, and an infinite horizon economy turned out to be the most tractable. Given the intricacies of strategic games in continuous time, a discrete approach turned out to be superior to a continuous time approach, especially since the increased tractability of the continuous time setting came at the cost of *ad hoc* assumptions needed about the profitability of firms after off the equilibrium path moves. By choosing a discrete state spaces with time invariant Markov transition processes we were able to use the powerful analytical tools available for such processes. Finally, having a continuum of industries, each of which containing a finite number of firms, allowed us to assume that each firm takes the market price of risk as given, while still generating aggregate general equilibrium effects of firms' strategic behavior.

The rest of the paper is organized as follows. In Section 2 we present the economic framework of the model. The equilibrium analysis of outcomes in each industry is presented in Section 3. Section 4 shows the existence of general equilibrium, and Section 5 analyzes how endogenous business cycles can arise. In Section 6, we study asset pricing implications of the model. Finally, concluding remarks are made in Section 7. All proofs are delegated to the Appendix.

2 Model Framework

2.1 Physical Environment

Consider an infinite horizon, discrete time, discrete state economy in which time is indexed by $t \in \mathbb{Z}_+$ and the time t state of the world is denoted by $s_t \in \{1, 2, \ldots S\}$.[1] Each period there is a transition between states which is governed by a Markov process with time invariant transition probabilities:

$$\mathbb{P}(s_{t+1} = j | s_t = i) = \Phi_{i,j}. \tag{1}$$

Here, $\Phi_{i,j}$ refers to the element on the ith row and jth column of the matrix $\Phi \in \mathbb{R}_+^{S\times S}$. We assume that Φ is irreducible and aperiodic, so that the process has a unique long-term stationary distribution.

2.1.1 Production

There is a continuum of industries, indexed by $z \in [0, 1]$, each consists of $N(z) \geq 1$ identical strategic firms that produce and sell a unique non-storable consumption good. The nature of the strategic environment is discussed in Section 2.2. The production technology for each good z at time t is linear in labor with stochastic productivity $A(z, t) = A_{s_t}(z)(1 + g)^t$. (That is, one unit of consumption good of z at time t in state s requires $\left[A_s(z)(1 + g)^t\right]^{-1}$ labor units.) Here, with some abuse of notation, $A_{s_t}(z)$ represents a state-dependent and sector-specific productivity component whereas $g \geq 0$ represents a common long-term productivity growth rate across all sectors. For tractability we assume that $A : S \times [0, 1] \to \mathbb{R}_{++}$ is a function that satisfies standard integrability conditions so that aggregation across industries is possible. Labor is supplied inelastically by a representative agent, who in each period divides her one unit of human capital across all the industries and earns a competitive wage, $w(t)$, in return.

[1] Here, $\mathbb{Z}_+ = \{0\} \cup \mathbb{N} = \{0, 1, \ldots\}$ is the set of non-negative integers. Also, we follow the standard convention that $\mathbb{R}_+$ is the set of nonnegative real numbers, whereas $\mathbb{R}_{++}$ is the set of strictly positive real numbers.

2.1.2 Preferences / Demand

The representative agent possesses CRRA preferences over aggregate consumption with risk aversion parameter γ and subjective discount factor $\hat{\delta}$, i.e.,:

$$U = \mathbb{E}\left[\sum_{t=0}^{\infty} \hat{\delta}^t \frac{C(t)^{1-\gamma}}{1-\gamma}\right], \tag{2}$$

where $C(t)$ represents the Dixit-Stiglitz *CES* consumption aggregator of goods (see Dixit and Stiglitz, 1977).[2]

$$C(t) = \left(\int_0^1 c(z,t)^{\frac{\theta-1}{\theta}} dz\right)^{\frac{\theta}{\theta-1}}. \tag{3}$$

The parameter $\theta > 1$ is the (constant) elasticity of substitution across goods. We note in passing that preferences with a more general state dependent utility specification are also covered by our specification.[3] The *CES* specification leads to standard *period-by-period* demand functions as a function of prices $p(z,t)$ and total income $y(t)$:[4]

$$c(z,t) = \frac{y(t)}{p(z,t)^{\theta} P(t)^{1-\theta}}, \tag{4}$$

where $P(t) \equiv \left(\int_0^1 p(z,t)^{1-\theta} dz\right)^{\frac{1}{1-\theta}}$ can be interpreted as the ideal price index (see Appendix A), total income $y(t)$ is derived from wages, and distribution of firm profits, $\pi(z,t)$, across all sectors z:

$$y(t) = w(t) + \int_0^1 \pi(z,t) dz, \tag{5}$$

$$\pi(z,t) = \left[p(z,t) - \frac{w(t)}{A(z,t)}\right] c(z,t). \tag{6}$$

Going forward, we will normalize the nominal price index $P(t)$ to 1.[5] Hence, income $y(t) = C(t)$, wages and profits are measured in units of aggregate consumption.

2.2 Strategic Environment

Within each industry z, $N(z)$ identical firms play a dynamic Bertrand pricing game with perfect public information as in Rotemberg and Saloner (1986). The timing of the stage game in each period, t, is as follows. First, the state, s_t is revealed. Then all firms $i \in \{1, 2, \ldots N(z)\}$ in industry z simultaneously announce their gross markup, $Q^{(i)}(z,t)$. For tractability, we express each

[2]See van Binsbergen (2007) or Ravn et al. (2006) for using *CES* preferences in a dynamic context.

[3]Consider the more general $\tilde{C}(t) = \left(\int_0^1 v_{s_t}(z) c(z,t)^{\frac{\theta-1}{\theta}} dz\right)^{\frac{\theta}{\theta-1}}$ as in Opp (2010). The state dependent "taste" function $v_s(z)$ can then easily be reduced to the case where $v_s(z) \equiv 1$, by transforming the productivity, $A_s(z) \mapsto v_s(z)^{(\theta-1)/\theta} A_s(z)$. Such a transformation can be interpreted as a numeraire change, where the amount of a unit of goods is redefined in each state. A state dependent taste function could, for example, represent an agent's higher utility of an umbrella in a rainy state than in a sunny state of the world.

[4]The demand functions $c(z,t)$ yield maximal $C(t)$ given an arbitrary price vector $p(z,t)$ and income $y(t)$. They are obtained via simple first-order conditions.

[5]This is without loss of generality, since the wage rate $w(t)$ is a free variable.

firm's strategy in terms of gross markups instead of prices, satisfying $p^{(i)}(z,t) = Q^{(i)}(z,t)\frac{w(t)}{A(z,t)}$. Consumers demand the product from the producer with the lowest markup. If all firms announce the same Q, total demand in sector z is evenly shared between all $N(z)$ firms. The firms then go out and hire workers to meet demand.

Each industry z coordinates on the *symmetric, subgame* perfect equilibrium outcome that exhibits a maximal degree of collusion, i.e., maximal industry profits. Due to symmetry, the equilibrium gross markup function of each firm i satisfies: $Q^{(i)}(z,t) = Q(z,t)$, with the associated industry price

$$p(z,t) = Q(z,t)\frac{w(t)}{A(z,t)}. \tag{7}$$

While the equilibrium outcome of this game is in general non-trivial (see Section 3.3), the two polar cases of a monopoly, i.e., $N(z) = 1$, and perfect competition provide useful bounds. If the industry is served by a monopolist, he maximizes industry profits (equation 6) subject to consumer demand (equation 4) which leads to an optimal markup of:

$$Q^m(z,t) = Q^m = \frac{\theta}{\theta - 1}.$$

If, on the other hand, $N(z)$ is infinite, then we expect prices to be set competitively. In this case, the markup is 1. If the number of firms is finite but greater than one, we expect equilibrium markups to be somewhere in between the competitive and monopolistic prices, i.e., $Q \in \left[1, \frac{\theta}{\theta-1}\right]$.[6]

3 Analysis

Before proceeding with our formal equilibrium analysis, it is convenient to transform our growing economy into a time-invariant economy in which outcomes only depend on time t through the state at time s_t. The resulting implications and other normalizations are presented in Section 3.1.

Our partial equilibrium analysis consists of two parts. First, for an arbitrary exogenous distribution of markups across industries, we characterize aggregate consumption, and show that it together with a measure of aggregate markups determines the efficiency losses in the economy (Section 3.2). Second, given the aggregate consumption and aggregate markup dynamics, we solve for the partial equilibrium outcome of one sector z in the economy, i.e., the optimal state-contingent markups (Section 3.3).

3.1 Preliminaries

We focus on equilibria which—except for the constant growth rate g—are time invariant in that outcomes are the same at t_1 and t_2 if the states are the same, i.e., if $s_{t_1} = s_{t_2}$. In other words,

[6] Although the condition that $Q \in \left[1, \frac{\theta}{\theta-1}\right]$ arises naturally, for the results in this section it is not crucial, and we need only assume the weaker condition that Q is strictly positive for all z and s.

we study equilibria on the form:

$$\begin{aligned}
C(t) &= (1+g)^t C_{s_t}, \\
y(t) &= (1+g)^t y_{s_t}, \\
w(t) &= (1+g)^t w_{s_t}, \\
\pi(z,t) &= (1+g)^t \pi_{s_t}(z), \\
c(z,t) &= (1+g)^t c_{s_t}(z),
\end{aligned}$$

where variables on the right hand side are growth-normalized, time invariant, variables which only depend on the state, s_t. In such an economy we immediately obtain that markups are time-invariant

$$Q(z,t) = Q_{s_t}.$$

Thus, our following analysis will suppress the time t dependence of all real variables and use the state-dependent form. We note that this does not in any way restrict the off-equilibrium path behavior, it is a restriction on the on-equilibrium path behavior of agents, and the corresponding equilibrium outcomes.

It follows from a standard transformation, using the utility representation (equation 2), that growth-normalized variables can be determined by solving the model for a non-growing economy with a growth-adjusted personal discount rate, i.e., with

$$\delta \stackrel{\text{def}}{=} (1+g)^{1-\gamma}\hat{\delta}. \tag{8}$$

Intuitively, the representative agent's tradeoff between consumption in different times and states is affected in identical ways by changes in the growth rate and the personal discount rate. We will use the formulation with δ going forward.

For ease of exposition, we decompose productivity shocks $A_s(z)$ into the functions $\alpha_s(z)$ and $\bar{A}_s$ where $\alpha : S \times [0,1]$ and the vector $\bar{A} \in \mathbb{R}^S_+$. Specifically,

$$\alpha_s(z) \equiv \frac{A_s(z)^{\theta-1}}{\int_0^1 A_s(z)^{\theta-1}dz} = \left(\frac{A_s(z)}{\bar{A}_s}\right)^{\theta-1}, \qquad \text{where} \tag{9}$$

$$\bar{A}_s \equiv \left[\int_0^1 A_s(z)^{\theta-1}dz\right]^{\frac{1}{\theta-1}}. \tag{10}$$

Here, $\bar{A}$ represents the average productivity shock to the economy and $\alpha_s(z)$ captures the industry productivity shock relative to the economy. In other words, changes in $\alpha(z)$ across states are *idiosyncratic* shocks to individual industries, whereas changes in $\bar{A}$ are *systematic* shocks. We can also view $\alpha(z)$ as an S-vector, $\alpha(z) \in \mathbb{R}^S$.

As a result of the normalization, the average relative industry state is equal to one, i.e., $\int_0^1 \alpha_s(z)dz = 1$. Now instead of specifying A, we can equivalently specify the function of idiosyncratic shocks, α, and the vector of systematic shocks, $\bar{A} \in \mathbb{R}^S_{++}$. Given the previous argument, the exogenous variables in the economy can then be represented by the tuple $\mathcal{E} = (\alpha, \bar{A}, N, \Phi, \theta, \gamma, \delta)$.

3.2 Aggregate Consumption

Aggregate consumption is an important endogenous variable. As outlined above, we will first treat the outcome of the strategic game for each industry and each state as exogenously given, as summarized by the gross markup functions $Q_s(z)$. Together with the functions, $\alpha_s(z)$ and $\bar{A}_s$, the real outcome in the economy or the consumer's consumption bundle is completely determined, state-by-state. We will use aggregate consumption in two ways. First, as a measure of inefficiency and second to price claims on financial assets.

3.2.1 Pareto Efficiency and Aggregate Markups

This section reveals that our economy can exhibit Pareto inefficient outcomes which distinguishes our paper from a standard real business cycle model.[7] Our inefficiency is generated by firms' value-maximizing, strategic price setting behavior and the attendant distortion in labor allocation.

For ease of exposition, we introduce two sufficient statistics of the cross-sectional markup distributions for the macro-economy in each state s:

$$\bar{Q}_s = M_{1-\theta}\left(Q_s\right), \tag{11}$$

$$e_s = \left(\frac{M_{-\theta}\left(Q_s\right)}{M_{1-\theta}\left(Q_s\right)}\right)^{\theta} \leq 1. \tag{12}$$

where $M_p\left(Q_s\right) = \left(\int \alpha_s(z) Q_s\left(z\right)^p dz\right)^{\frac{1}{p}}$ refers to the p-th order cross-sectional power mean of $Q_s\left(z\right)$.[8] Both sufficient statistics capture distinct elements of the cross-sectional markup distribution. The variable $\bar{Q}_s$ captures the notion of aggregate market power, i.e., an appropriate average markup across industries. The variable e_s captures the (inverse of) heterogeneity of markups across industries. By Jensen's inequality, e_s is bounded above by one (obtained when all industries charge the same markup) and is decreasing in the heterogeneity of markups.[9] It can also be interpreted as a measure of relative production efficiency, as the following proposition reveals.

Proposition 1. *Given the functions Q_s, α_s and $\bar{A}_s$, aggregate consumption, C_s, real income y_s, in state s are given by:*

$$C_s = y_s = \bar{A}_s e_s. \tag{13}$$

The fraction of real *income that is derived from labor income is given by:*

$$\omega_s = \frac{1}{e_s \bar{Q}_s}. \tag{14}$$

Real *firm profits in sector z are:*

$$\pi_s\left(z\right) = C_s \alpha_s\left(z\right) \bar{Q}_s^{\theta-1} \frac{Q_s\left(z\right) - 1}{Q_s\left(z\right)^{\theta}}. \tag{15}$$

The outcome in state s is Pareto efficient if $Q_s(z) \equiv k_s$ for all z, so that $e_s = 1$.

---

[7] In papers such as Kydland and Prescott (1982) and Long and Possner (1983), the outcome is efficient.

[8] Notice that by construction $\int_0^1 \alpha_s(z)dz = 1$, so we interpret α as a weighting measure where each industry obtains a weight according to its relative productivity.

[9] This follows from the fact that $M_p\left(\tilde{x}\right) > M_q\left(\tilde{x}\right)$ for any non-degenerate random variable $\tilde{x}$ as long as $p > q$.

From equation 13, aggregate consumption only depends on the aggregate shock $\bar{A}_s$ and the heterogeneity of markups embedded in e_s. Since $e_s \leq 1$, the upper bound of aggregate consumption, i.e., potential output, is given by the aggregate shock $\bar{A}_s$. If all industries in the economy are perfectly competitive, so that $Q_s(z) \equiv 1$, then it is easy to see that the outcome is Pareto efficient (i.e., $C_s = \bar{A}_s$). On reflection, it follows that any economy in which markups across industries are the same in each state (i.e., $Q_s(z) \equiv k_s$ for all z and s) will also be Pareto efficient. A special case of this is an economy in which all industries are monopolized as $Q_s^m(z) = \frac{\theta}{\theta-1}$ for all industries and states. We note that Equation 15 provides a bijection, $\pi_s \leftrightarrow Q_s$, where $1 \leq Q_s \leq \frac{\theta}{\theta-1}$, $0 \leq \pi \leq \zeta C_s \alpha_s(z) \bar{Q}_s^{\theta-1}$, and $\zeta = \frac{(\theta-1)^{\theta-1}}{\theta^\theta}$ is a constant.

The intuition for this result is simple. If markups are constant across industries, i.e., $e_s = 1$, relative goods prices are not distorted. Compared to perfect competition, however, the fraction of income derived from labor income will be lower, i.e., $\omega = \frac{\theta-1}{\theta}$ in the monopoly case, since firm profits derived from monopoly rents also accrue back to the representative agent. This feedback of monopoly profits into the budget constraint causes real demand for each good to be unaffected. That is, the labor allocation to industries is not distorted and aggregate consumption is therefore maximal.

We note that the efficiency of the completely monopolized outcome depends on our assumption that labor is the only factor input in the production function, which together with the fact that labor is always fully utilized implies that the relative allocation of labor across industries is what matters. In the fully monopolized case, markups are the same so no relative distortions exist across industries, hence the result. In this respect, the result is model dependent. However, the main points — which will also be valid in a more general setting — are that higher markups in some industries, in general equilibrium, do not necessarily lead to lower efficiency, and that the degree of markup *variation* across industries may be especially important in determining efficiency losses.

It follows from this discussion that the economically interesting case is one in which some industries have high markups and some low. In such economies, the distortions resulting from distorted allocation of labor, or alternatively the welfare costs, are the highest. Figure 1 illustrates the range of feasible welfare losses under the natural restriction that firms charge markups between 1 and $\frac{\theta}{\theta-1}$.

3.2.2 Valuation

To value claims, we assume that there is a complete market of Arrow-Debreu securities in zero net supply traded, in addition to the stocks of the firms. As a result, the unique one period stochastic discount factor of the time-invariant economy, SDF, satisfies:

$$SDF_{t+1} = \delta \left(\frac{C_{s_{t+1}}}{C_{s_t}} \right)^{-\gamma}. \tag{16}$$

This is the textbook stochastic discount factor for power utility, the only exception being that C_{s_t} now represents a properly defined consumption bundle (as opposed to a single consumption good).

For time-invariant economies, since time-t profits of a firm depend only on the state, s, the information about the firm's future profits can be summarized in an S-vector, π, where π_s is

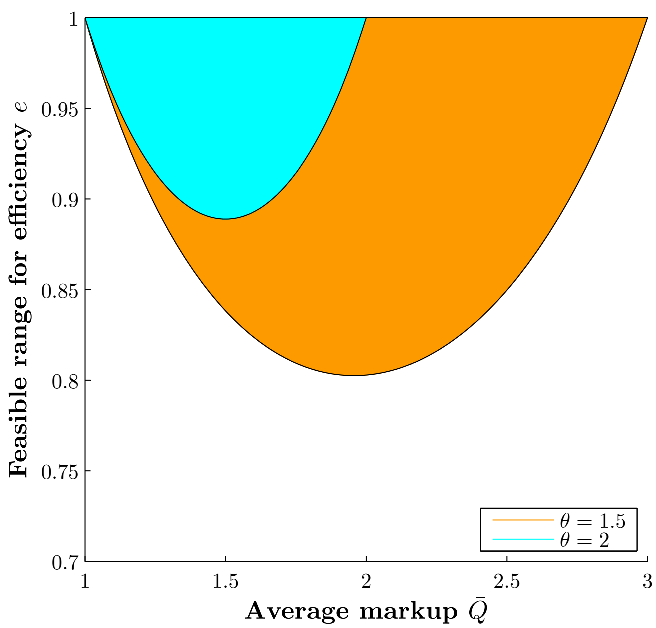

Figure 1. This graph plots the feasible range for efficiency e given an average markup of $\bar{Q}$ in the economy. In the monopolistic economy, i.e., when $\bar{Q} = \frac{\theta}{\theta-1}$, and the competitive economy, i.e., $\bar{Q} = 1$, there are no welfare losses. In the intermediate region, potentially large welfare losses may occur if some industries charge high markups and some low markups. These potential welfare losses are larger for smaller θ, i.e., for higher monopoly markups of $\frac{\theta}{\theta-1}$.

the profit in state s. We can then define the ex-dividend value vector V as an S-vector where the element V_s represents the the value of the firm in state s

$$V = \delta \Lambda_m^{-1} \Phi \Lambda_m (\pi + V). \tag{17}$$

Here, Λ_m is a diagonal matrix, with $m_s = C_s^{-\gamma}$ as its sth diagonal element. This argument leads to the following convenient formula for the value vector of the firm:

$$V = \Theta \pi, \tag{18}$$

where

$$\Theta = \Lambda_m^{-1} (I - \delta \Phi)^{-1} \Lambda_m - I, \tag{19}$$

and I is the $S \times S$ identity matrix. The valuation operator Θ has strictly positive elements. This simply represents the fact that higher profits in some state s increase the present value of future profits, $V_{s'}$, in all states $s' = 1, \ldots, S$.[10]

3.3 Industry equilibrium

Understanding strategic price setting behavior in one industry is the first step towards endogenizing Q. We therefore characterize, as a function of industry and aggregate characteristics,

[10]Recall that Φ is irreducible, so each state will be reached with positive probability, regardless of the initial state.

when firms in a specific industry behave competitively, when a monopolistic outcome can be sustained, and when the outcome is neither of these extremes. Observe that each industry is small compared with the aggregate economy, and that firms in industry z take the dynamics of all other industries as exogenously given, i.e., they take Q as exogenously defined for all $z' \neq z$. This is rational, since, as a result of Proposition 1, aggregate consumption C and average markups $\bar{Q}$ are not affected by an individual industry's behavior, and neither is then the pricing kernel. Thus, the relevant aggregate dynamics of the macro-economy for the outcome in any given industry are characterized by a $S \times 2$ matrix consisting of the vectors C and $\bar{Q}$.

As we have already observed, if an industry is monopolized, the outcome markup is $Q^m = \frac{\theta}{\theta-1}$ and when it is perfectly competitive the markup is $Q^c = 1$. Of course, in the latter case profits are zero, whereas in the former, monopolist profits, π_s^m are

$$\pi_s^m(z) = \zeta \bar{Q}_s^{\theta-1} C_s \alpha_s(z). \tag{20}$$

Here, since C_s and $\bar{Q}_s$ are macro variables they are systematic and affect profits positively in all industries. Real profits in a particular industry depend positively on the aggregate market power because goods are substitutable. The sector specific component of profits is given by the idiosyncratic productivity shock $\alpha_s(z)$.

Following Abreu (1988), we are interested in industry equilibria that generate the highest industry profits sustainable by credible threats. We restrict attention to *symmetric*, *pure strategy subgame perfect* equilibria of the infinitely repeated stage game described in Section 2.2. Firms condition their action at time t on the entire history of past actions of industry z and states up to time t. The relevant history of each industry z, h_t is defined as the entire sequence of markups, states, and aggregate variables:

$$h_t = \left\{ \left\{ Q^{(i)}(z,\tau) \right\}_{i=1}^{N(z)}, s_\tau, \bar{Q}_{s_\tau}, C_{s_\tau} \right\}_{\tau=0}^{t}, \tag{21}$$

with h_0 representing the empty history. Thus, a time-t, industry-z strategy for firm i is a mapping from $h_{t-1} \times S$ to a chosen markup, Q_τ^i, $f : h_{t-1} \times S \to R_{++} \in R_{++}^{h_{t-1} \times S}$. Here, the second parameter, $s \in S$, represents time t information about the state, which is available for the firm. A strategy for firm i is a sequence of time-τ strategies, $\left\{ f_\tau^i \right\}_{\tau=0}^{\infty}$.

The entire set of subgame perfect equilibria can be enforced with the threat of the worst possible subgame perfect equilibrium. In this case, the most severe punishment is given by the perfectly competitive outcome, i.e., zero profits forever after a deviation.

Any subgame perfect equilibrium must satisfy the following incentive constraints for each state s,

$$\frac{\pi_s(z) + V_s(z)}{N(z)} \geq \pi_s(z). \tag{22}$$

That is, the share of discounted present value of profits under collusion, $\frac{\pi_s + V_s}{N}$, must be greater or equal to the best-possible one period deviation of capturing the entire industry demand π_s and zero profits thereafter.[11]

[11] We are implicitly assuming that firms can coordinate within an industry to achieve this best outcome with this equilibrium selection mechanism. This trivially rules out any outcomes where markups are higher than $\frac{\theta}{\theta-1}$, and outcomes where markups are lower than necessary. We do *not*, however, assume that firms can coordinate across industries, since in a large economy there are many industries and global coordination therefore typically is not possible.

In the maximum profit equilibria, in each state, s, firms in an industry choose the vector of state contingent markups to maximize the value function, $V_s(z)$, given the value in each of the other states of the world, $V_{-s}(z)$, subject to incentive compatibility (equation 22),

$$V_s(z) = \arg\max_{Q_s} : V_s(z) \, | V_{-s}(z), \tag{23}$$

for all s. Here, Q_s maps to V_s via (15,18). Obviously, the case $N = 1$ is trivial: It leads to profits of π^m. We therefore focus on the case when $N \geq 2$.

In principle (23) is a complex optimization problem. First, strategic firms need to solve a state dependent infinite horizon state problem, second the optimal markup is a highly non-linear function of profits. However, within our model's setting, finding the solution is actually quite straightforward. First, we note that the dynamic equilibrium can be viewed as a linear programming problem in which firms choose profits instead of prices, replacing Q_s in (23) with π_s. Second, the specific form of this corresponding linear programming problem makes it clear that the solution is the same for each state, and the optimization therefore collapses to a static, state independent, linear programming problem.

To see this, we first consider a relaxed optimization problem starting from an arbitrary state j in which we ignore the additional consistency requirements imposed by value-maximizing behavior in other states $s \neq j$, i.e.,

$$\pi^j(z) = \arg\max_{\hat{\pi}(z)} \; \iota_j^T \Theta \hat{\pi}(z), \quad \text{s.t.}, \tag{24}$$

$$\hat{\pi}(z) \leq \pi^m(z), \tag{25}$$

$$0 \leq (\Theta - (N(z) - 1) I) \hat{\pi}(z), \tag{26}$$

where ι_j refers to an S-vector with zeros, except for the jth element which is equal to unity, $\iota_j = (\underbrace{0, \ldots, 0, 1}_{j}, \underbrace{0, \ldots, 0}_{S-j})^T$. Feasible profits are bounded above by monopoly profits which is captured by equation 25. The vector of IC constraints following from equation 22 is captured by equation 26.

Lemma 1. *The maximizer of the relaxed program, $\pi^j(z)$, described in equations 24 - 26 is independent of the initial state j, i.e.,*

$$\pi^j(z) = \pi(z), \tag{27}$$

for some $\pi(z) \in \mathbb{R}_+^S$. Further, for each element, s, either constraint 25 or 26 binds in the solution, $\pi_s(z)$.

This technical Lemma implies that the consistency requirements are automatically satisfied leading to the following Proposition.

Proposition 2. *Given aggregate consumption C and the average markup $\bar{Q}$, the industry equilibrium outcome $\pi(z)$ (or equivalently $Q(z)$) is unique. Equilibrium profits in state s are either given by monopoly profits, $\pi_s^m(z)$, or the IC constraint in state s binds, i.e. $\pi_s(z) = \frac{\iota_s^T \Theta \pi(z)}{N(z)-1}$.*

Going forward, it will be important to understand when the incentive constraint binds. This is because, as we have observed, Pareto inefficiencies arise if markups differ across industries. To measure the "tightness" of the monopolistic incentive constraint, we introduce the "tightness" vector, $\Gamma(z)$, with element s denoting the s-state ratio of the present value of industry profits under monopoly markups to monopoly profits:

$$\Gamma_s(z) = \frac{\pi_s^m(z) + V_s^m(z)}{\pi_s^m(z)} = 1 + \frac{V_s^m(z)}{\pi_s^m(z)}. \tag{28}$$

If $\Gamma_{s_1} > \Gamma_{s_2}$ the incentive to deviate in state s_1 is smaller than in state s_2, i.e., the present value of collusion is high relative to current period profits.

Lemma 2. *The tightness vector satisfies:*

$$\Gamma(z) = (\Lambda_{\kappa(z)}^{-1}(I - \delta\Phi)^{-1}\Lambda_{\kappa(z)})\mathbf{1}, \tag{29}$$

where $\Lambda_{\kappa(z)} = diag(\kappa(z))$, *and the vector* $\kappa(z)$ *has elements:*

$$\kappa_s(z) = \pi_s^m(z)\, m_s = \zeta \bar{Q}_s^{\theta-1} C_s^{1-\gamma} \alpha_s(z).$$

The variable κ_s captures an important determinant of the incentive to cheat in a certain state, Γ_s. It consists of the state component of the industry profit, $\pi_s^m(z)$, weighted by marginal utility in state s, $m_s = C_s^{-\gamma}$. We also define the minimum, $\underline{\kappa}(z) = \min_s \kappa_s(z)$.

Using the definition of the tightness vector, we are now able to derive closed-form expressions for the threshold number of players that leads to perfect competition and the monopoly outcome, respectively. Intuitively, for few enough players $N(z) \le N^m(z)$, the monopoly outcome is sustainable in all states, while too many firms in one industry, $N(z) \le N^c$, generates the competitive outcome in all states. In between, markups may vary across states. This intuition is formalized in the following proposition.

Proposition 3. *Given aggregate consumption C and the average markup $\bar{Q}$, equilibrium profits in state s, $\pi_s(z)$, satisfy:*

$$\begin{array}{lll} \pi_s(z) = \pi_s^m(z) & \textit{for} & N(z) \le N^m(z), \\ \pi_s(z) \in (\kappa(z) C_s^\gamma, \pi_s^m(z)] & \textit{for} & N(z) \in (N^m(z), N^c), \\ \pi_s(z) = C_s^\gamma \underline{\kappa}(z) & \textit{for} & N(z) = N^c, \\ \pi_s(z) = 0 & \textit{for} & N(z) > N^c. \end{array}$$

where the respective threshold values satisfy $N^m(z) \stackrel{\text{def}}{=} \min_s(\Gamma_s(z))$ *and* $N^c \stackrel{\text{def}}{=} \frac{1}{1-\delta}$.

The different regions are best shown in example. Assume that aggregate consumption satisfies $C = (1,2,4)^T$, that aggregate markups are competitive in all states, $\bar{Q} = (1,1,1)^T$, and that $\alpha(z) = \left(\frac{1}{2}, \frac{3}{4}, 1\right)^T$. Moreover, assume preference parameters of $\delta = 8/9$, $\gamma = 2$, and $\theta = 2$. Is is easy to show that the tightness vector in this example satisfies:

$$\Gamma(z) = (7, 9, 13)^T \tag{30}$$

Thus, monopoly markups are sustainable for $N(z) \leq N_M(z) = 7$.[12] Given δ, the number of firms necessary to induce the competitive outcome is $N_C = 9$. Figure 2 plots state-contingent profits in the left panel and the corresponding state-contingent markups as a function of the number of firms, confirming the four cases in Proposition 3.

Note that this example features pro-cyclical markups, i.e., $Q_1(z) \leq Q_2(z) \leq Q_3(z)$. While profits are unusually high in the good state of the world (increasing the incentive to cheat), this effect is overwhelmed by the valuation effect, i.e., future profits are discounted at a lower rate in times of low marginal utility. In contrast, in the worst state of the world the incentive to cheat is exacerbated by high marginal utility.

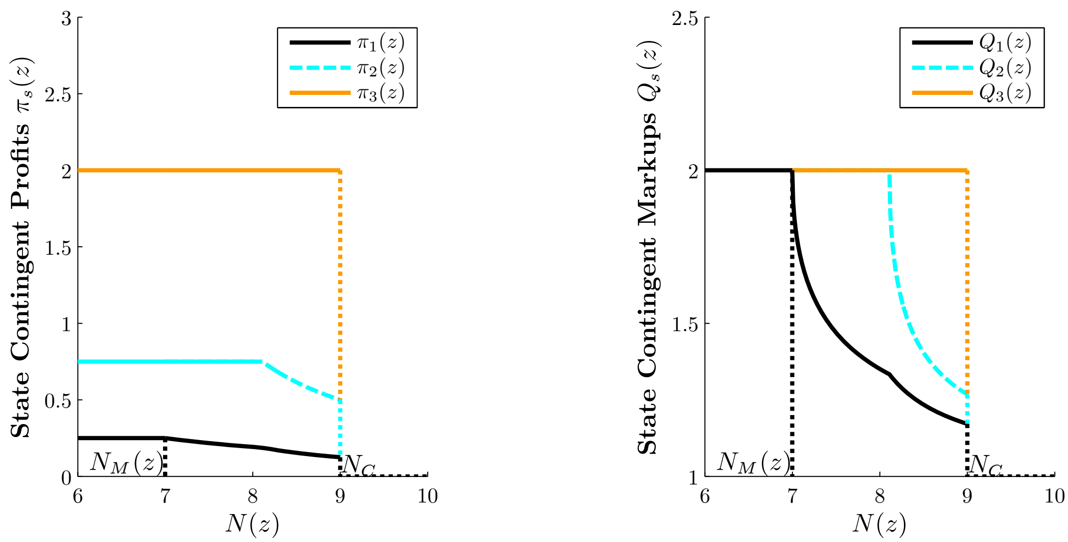

Figure 2. This graph plots the state contingent profits and markups of one particular industry given aggregate consumption of $C = (1, 2, 4)^T$, aggregate markups of $\bar{Q} = (1, 1, 1)^T$, and the relative industry state of $\alpha(z) = (2, 3, 4)^T$. If there are fewer than 7 firms in the industry, monopoly markups are sustainable in all states. Increasing the number of firms further cause the incentive constraint in state 1 to bind first, then in state 2 and finally, at $N_C = 9$, all markups collapse discontinuously to the competitive outcome, i.e., 1.

Given the generality of our setup it is quite surprising how much structure we are able to put on the industry equilibrium outcome. Intuitively, the threshold number of firms that allows the monopoly outcome is directly linked to $\Gamma(z)$. It is determined by the state in which the incentive to deviate is the highest, i.e., the state in which $\Gamma(z)$ attains its minimum. Secondly, the maximum number of players beyond which collusion completely breaks down is simply given by $N^c \stackrel{\text{def}}{=} \frac{1}{1-\delta}$, i.e., it only depends on the growth adjusted discount rate. Quite surprisingly, the threshold value is independent of industry characteristics as captured by $\alpha(z)$ and aggregate properties such as aggregate consumption C or the average markup $\bar{Q}$. At N^c, the incentive constraint is characterized by the indifference condition of a risk-neutral firm that compares the shared perpetuity value under collusion, $\frac{\pi^*(z)}{1-\delta}\frac{1}{N^c}$, and the best possible one-period deviation, $\pi^*(z)$. Moreover, we are able to derive an analytical formula for the profits of any industry z

[12]Recall that aggregate markups are competitive even if a zero measure of industries are non-competitive. Thus, there is no inconsistency in having one non-competitive industry in an economy that in aggregate is competitive.

with N^c firms.

What remains is to characterize the solution for industries with $N^m(z) < N(z) < N^c$ firms. For a special case, this region is empty, i.e., if $N^m(z) = N^c$.

Lemma 3. *If $\kappa_s(z) = k$ for all s and some arbitrary constant k, the threshold value for the monopoly outcome is given by N^c, i.e.: $N^m(z) = N^c$.*

In such industries markups are never state dependent (regardless of the number of firms in the industry), since they are neither state dependent in the monopolistic case, nor in the competitive case. Except for this knife-edge case, the region between $N^m(z)$ and N^c is non-empty, and represents the economically most interesting region, since it gives rise to state-contingent markups.

Proposition 4. *For an industry in which $N^m(z) < N(z) < N^c$,*

1. *There will be at least one state in which monopolistic profits are obtained, $\pi_s(z) = \pi_s^m(z)$ for some s.*
2. *Equilibrium profits, $\pi_s(z)$, are nonincreasing in $N(z)$ for each s, as are markups.*
3. *Equilibrium profits, $\pi_s(z)$, are nondecreasing in $\alpha_{s'}(z)$, for each s, s', as are markups.*
4. *Equilibrium profits and markups depend continuously on all parameters (N, C, $\bar{Q}$, Φ, α, and $\bar{A}$).*

It is straightforward to verify properties 1, 2, and 4 in Figure 2. Thus, given that the aggregate variables of the economy are known, the qualitative behavior of markups in different states of the world in a specific industry is well understood.

4 General Equilibrium

We show the existence of general equilibrium in which firms in each industry choose optimal markups given the (optimal) markups chosen by firms in all other industries. Recall that the economy's environment is characterized by the tuple $\mathcal{E}$, i.e., by the real variables $\alpha : S \times [0,1] \to \mathbb{R}_+$, $N : [0,1] \to \mathbb{N}$, $g \geq 0$, $\bar{A} \in \mathbb{R}^S_{++}$, the irreducible aperiodic stochastic matrix, $\Phi \in \mathbb{R}^{S \times S}_{++}$, and the preference parameters, γ, θ, and $\hat{\delta}$. We note that a given equilibrium is completely characterized by the markup function, $Q : S \times [0,1] \to \left[1, \frac{\theta}{\theta - 1}\right]$, together with $\mathcal{E}$, since all other real and financial variables can be calculated from Q using (7) and (11-15). This motivates the following

Definition 1. *General Equilibrium in economy $\mathcal{E}$ is given by a markup function $Q : S \times [0,1] \to \left[1, \frac{\theta}{\theta-1}\right]$ for which,*

1. *$\bar{Q}$ and C are defined by Equations 11 and 13,*

2. *For all z, $Q(z)$ is the solution to the maximization problem given by Equations 24-26, where $\pi^m(z)$ in the optimization problem is given by Equation 20.*

We note that the existence and uniqueness of the second part of the definition is guaranteed by Proposition 2, industry by industry, i.e., given $\bar{Q}$ and C there is a unique optimal markup function. It is a priori unclear, however, whether there exists a general equilibrium, i.e., whether both parts can be solved simultaneously. In other words, both the mappings, $Q \mapsto (\bar{Q}, C)$ (part 1) and $(\bar{Q}, C) \mapsto Q'$ (part 2) are well defined, but it is unclear whether Q can be chosen such that the second step maps to the same markup function that was used in the first step, i.e., such that $Q' = Q$.

It turns out that we are able to prove the existence of equilibrium under very general conditions. Specifically, we assume that the functions N and α are Lebesgue measurable functions, and impose the following technical condition:

Condition 1. *For all s, for almost all z, $c_0 \leq \alpha_s(z) \leq c_1$ for constants, $0 < c_0 \leq c_1 < \infty$.*

Before showing existence, we discuss some invariance results which will be helpful in the proof. We first note that the following result follows immediately from Proposition 2:

Lemma 4. *In any general equilibrium, any two industries with the same N and α have the same markups, Q, and profits, π.*

Also, we observe that it is only the distributional properties of N and α that are important for the aggregate characteristics of an equilibrium. This should come as no surprise given that the aggregate variables important for industry equilibrium only depend on the distributions. To be specific, we define the (cumulative) distribution function $F : \mathbb{N} \times [c, C]^S \to [0, 1]$, where $F(n, s_1, \dots, s_S) = \lambda(\{(z : N(z) \leq n \wedge \alpha_1(z) \leq s_1 \wedge \dots \wedge \alpha_S(z) \leq s_S\})$, and λ denotes Lebesgue measure. Thus, $F(n, \alpha_1, \dots, \alpha_S)$ denotes the fraction of industries with number of firms less than or equal to n, and productivities $\alpha_s(z) \leq \alpha_s$ for all s. We then have

Lemma 5. *Given two economies, 1 and 2, that are identical except for that their functions determining number of firms and productivity, N and α, differ. Assume that the economies have the same distribution function, F. Then they have have the same equilibria in the sense that for each equilibrium in the first economy, there is an equilibrium in the second, such that any two industries, z and z' in the first and second economy, respectively, for which $N^1(z) = N^2(z')$ and $\alpha_s^1(z) = \alpha_s^2(z')$ for all s, have the same industry markups in each state of the world, $Q_s^1(z) = Q_s^2(z')$ for all s.*

We now have the following general result:

Proposition 5. *General equilibrium exists in any economy that satisfies Condition 1.*

Thus, only the technical conditions of integrability and boundedness of productivity functions across industries is needed to ensure the existence of equilibrium.

The proof of Proposition 5 uses Schauder's fixed point theorem. Specifically, the continuity results from the previous section implies that the mapping $Q \mapsto (\bar{Q}, C) \mapsto Q'$ is continuous (technically, in the function space L^1). It is further shown in the proof of the Proposition 5 that the space of markup functions is compact and convex, which via Schauder's theorem then guarantees the existence of a fixed point, i.e., an equilibrium. We note that Proposition 5 makes no claim about equilibrium uniqueness — a subject that will be explored further in the next section.

5 Endogenous Business Cycles

In this section we analyze general equilibrium. The analysis is qualitative, without any attempt toward a real world calibration. Such a calibration would be highly interesting — and potentially feasible given the our flexible set-up — but is outside of the scope of this paper. Our main objective is to show — with a sequence of examples — how strategic competition may endogenously amplify and even generate business cycles.

5.1 Business Cycle Amplification

Previous asset pricing literature has mainly studied the effects of strategic interaction in a partial equilibrium setting. In general equilibrium the decisions of firms in one part of the economy, via the influence they have on aggregate consumption, affect the pricing kernel and thereby the decisions of all other firms in the economy.

We introduce the following measure of business cycle amplification through strategic interaction:

Definition 2. *The* Oligopolistic Business Cycle Amplifier *(OBCA) is defined as*

$$OBCA = \frac{\sigma_C}{\sigma_{\bar{A}}}. \tag{31}$$

Here, σ_C and $\sigma_{\bar{A}}$ are the unconditional standard deviations of aggregate consumption and aggregate productivity, respectively.

Recall that in the first-best competitive outcome, $C \equiv \bar{A}$, leading to an OBCA equal to one. Business cycle fluctuations in this case are purely technology-driven. An OBCA greater than one thus tells us that equilibrium business cycle fluctuations, which are also affected by firms' strategic behavior, are larger than what is motivated by technological shocks. Similarly, an OBCA less than one implies that strategic behavior in equilibrium dampens business cycle fluctuations.[13]

We note that since $C = e\bar{A}$ (see Equation 13), the variation of C will be especially high when e and $\bar{A}$ are positively co-dependent, leading to an all else equal higher OBCA. Further, since e_s is determined by the cross-sectional dispersion of markups across industries in state s

[13] Our OBCA is defined in absolute terms. A "relative" OBCA, a ROBCA, is defined as $ROBCA = \frac{\sigma_C/E[C]}{\sigma_{\bar{A}}/E[\bar{A}]}$. Since $E[C] \leq E[\bar{A}]$, it trivially follows that $OBCA \leq ROBCA$, so business cycle amplifications are at least as severe in relative terms.

(see Equation 12), such that when the dispersion is low then e_s is high (close to one), whereas when the dispersion is high, e_s is low (close to zero), it follows that oligopolistic business cycle amplifications should be especially large in economies in which the cross sectional dispersion of markups is high in low-productivity states.

Consider the economy described in Table 1, with three distinct types of industries, I_1, I_2 and I_3, and $S = 2$ states.

Type, j	I_j	N	A_1	A_2	α_1	α_2
1	$z \in [0, 0.02)$	19	0.25	1	0.8728	1
2	$z \in [0.02, 0.81)$	19	1	1	1.0026	1
3	$z \in [0.81, 1]$	1	1	1	1.0026	1
$\bar{A}$					$\bar{A}_1 = 0.974$	$\bar{A}_2 = 1$

$$\Phi = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$$

$$\gamma = 6, \qquad \theta = 1.1, \qquad \delta = 0.95.$$

Table 1. Economy with three industries and two states.

Here, all industries with $z \in I_j$ belongs to industry type j. With a slight abuse of terminology, we will call the I_j sets "industries," although each set represents many identical industries. Thus, there is one very small industry (I_1), one large industry (I_2) and one medium-sized industry (I_3). The first two industries have many firms, $N = 19$, but they will still not be perfectly competitive, since $N^c = \frac{1}{1-\delta} = 20$. The third industry is monopolistic, $N = 1$.

Columns 4 and 5 in Table 1 describe the absolute productivity shocks, A, in the two states. We see that only the very small first industry experiences any variation in productivity across the two states. The aggregate variation in productivity will therefore be small. In columns 6 and 7, we show the decomposition of the absolute productivity shocks into relative and aggregate components, α and $\bar{A}$ (see (9) and (10)). The effect on aggregate productivity of the first industry's shock is about 2.5%, since aggregate productivity is 0.974 in the low-productivity state and 1 in the high-productivity state. This would also be the aggregate consumption in the two states in an efficient outcome. Note that the shock to industry 1 also affects the relative productivity in industries 2 and 3, since α is normalized to sum to one across industries, state by state.

Before analyzing the equilibrium in this economy, it is instructive as a reference case to study the economy which is identical to that in Table 1, except for that $A_1 = 1$ in industry 1. This is thus an economy with no productivity shocks, neither idiosyncratic nor aggregate, and it follows that $\bar{A}_1 = \bar{A}_2 = 1$ and $\alpha_s(z) \equiv 1$ in this reference economy. One easily verifies that the monopolistic outcome, in which markups $Q \equiv \frac{\theta}{\theta-1} = 11$ are chosen by all firms in all states, is feasible in this case (this also follows as a consequence from Lemma 3, since $N \leq N^c$ in all industries), leading to the efficient outcome where $C_1 = \bar{A}_1 = 1$, $C_2 = \bar{A}_2 = 1$.

The situation is different for the economy given in Table 1. The fully monopolistic outcome is no longer feasible, because it does not satisfy the IC constraints for firms in industry 1.

Instead, an equilibrium is given by the following markups:

$$\begin{array}{l|ll} \text{Markups} & s=1 & s=2 \\ \hline Q(I_1) & 1.580 & 11 \\ Q(I_2) & 1.465 & 11 \\ Q(I_3) & 11 & 11 \\ & & \end{array} \tag{32}$$

leading to aggregate consumption

$$C_1 = 0.795, \qquad C_2 = 1.$$

Thus, the small productivity shock ($\approx$ 2.5%) leads to a significant decrease in equilibrium output ($\approx$ 20%) in state 1. The OBCA in this equilibrium is

$$OBCA = \frac{\sigma_C}{\sigma_{\bar{A}}} = \frac{0.103}{0.013} = 7.88,$$

so strategic interaction leads to an almost eight-fold amplification of the business cycle variation. The intuition for why this amplification occurs is exactly in line with our main theme in this paper, that technological shocks that are small in aggregate — in that they only affect a few industries — change the strategic behavior of firms in other industries through the effect they have on the pricing kernel.

This mechanism is explained in Figure 3, focusing on the behaviors of industries 1 and 2.[14] In the upper part of the figure, the reference economy with identical industries is shown, in which case monopolistic profits are feasible for both industries. In the lower part of the figure, the economy in Table 1 is shown. Line A shows the important IC constraint, given the pricing kernel in the monopolistic outcome. Monopolistic profits are indeed feasible in industry 1 (lower left figure), but infeasible in industry 2 (lower right figure). Thus, the lower productivity in industry 1, through its effect on the pricing kernel, affects the outcome in sector 2, which moves to line B. This in turn changes the pricing kernel even further, making monopolistic profits in industry 1 infeasible and further changing the outcome in industry 2, moving to lines C in the two industries, and generating further feedback effect. The ultimate effect of this mechanism is that the equilibrium moves to line D in the two figures, substantially different from monopolistic equilibrium in the reference economy.

The previous result highlight that although our model is based on a similar framework as real business cycle models (Kydland and Prescott (1982), Long and Possner (1983)), equilibrium dynamics may be quite different. Specifically, significant business cycle fluctuations may arise even when aggregate "technological" shocks are small. A recent strand of literature has aimed at explaining how technological shocks at the individual firm or industry level do not diversify out, but may affect aggregate productivity. Gabaix (2011) notes that if the distribution of firm size is heavy-tailed, firm-specific shocks may indeed affect aggregate productivity. Acemoglu et al. (2011), suggest that inter-sectoral input-output linkages between industries may lead to

[14]Industry 3 is always monopolistic. The reason that it is still important for the example is that substantial efficiency losses only occur when there is high variability in markups across sectors. If industry 3 was not present then the economy would always be close to efficient, since markups would be the same for the vast majority of industries in each state — almost identical to the markups charged in industry 2. In contrast, when industry 3 is present and industry 2 charges low markups, efficiency will be low.

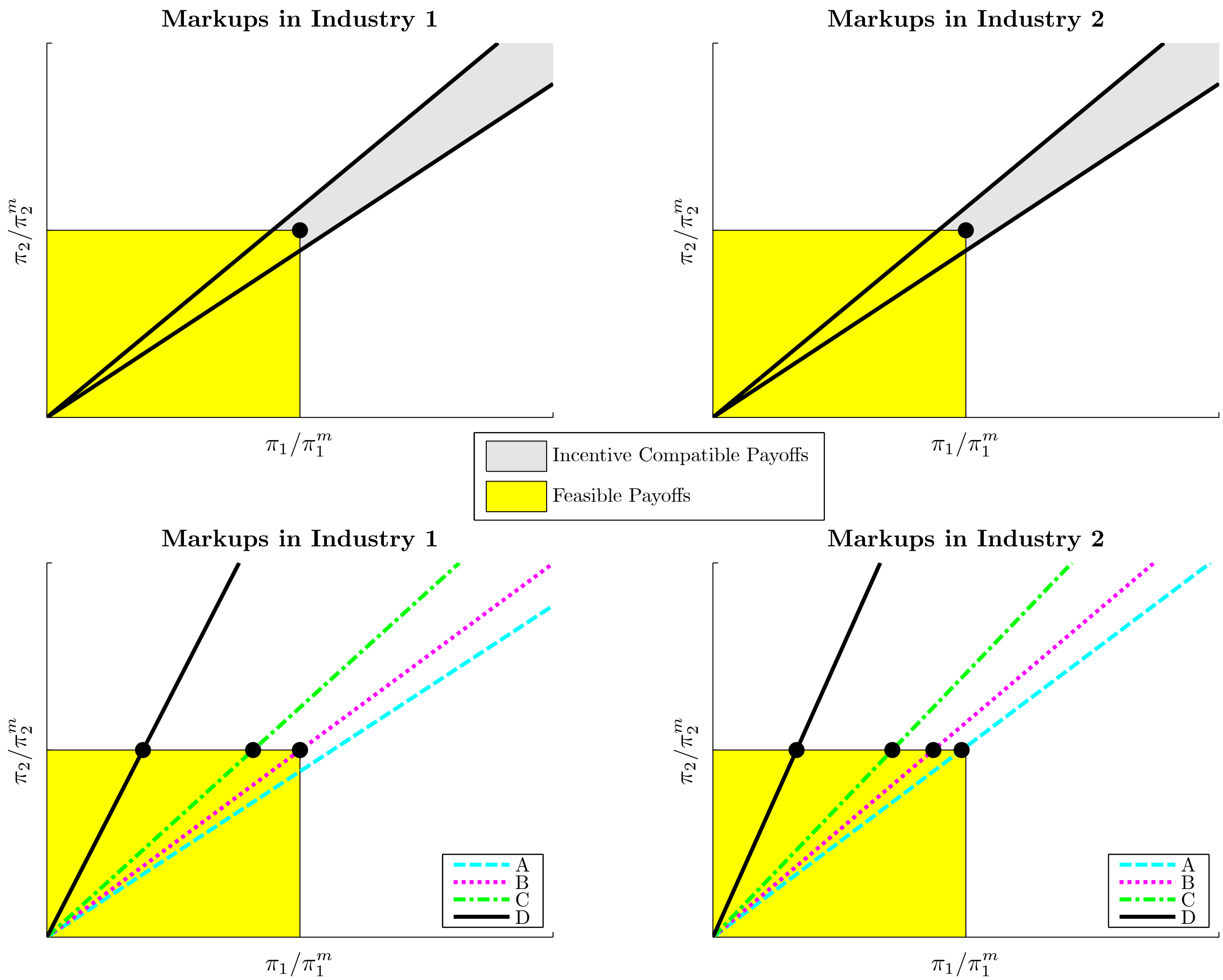

Figure 3. Profits in industry 1 (left) and 2 (right) in economies with identical industries (above) and economy given in Table 1 (below). Above: Monopolistic profits are feasible. Below: Monopolistic profits violate IC constraint of industry 1, in turn changing the IC constraints in industry 2. The resulting equilibrium is substantially different.

"cascades effects" where a shock in one industry spreads through the economy and thereby becomes an aggregate shock.

The mechanism in our model is quite different, more along the lines suggested in Jovanovic (1987), who shows that idiosyncratic shocks may not cancel out in strategic games with a large number of players. In our previous example, aggregate productivity is close to constant across states, but because it varies at the sectoral level, the strategic behavior of firms leads to aggregate shocks in equilibrium. We believe that this provides an important mechanism for understanding the sources of aggregate fluctuations in the economy.

5.2 Comparative Statics

The equilibrium outcome may be very sensitive to small changes in some parameter values, whereas it is remarkably stable in other aspects. The results together suggests that cross economy (e.g., cross country) comparisons need to be carefully designed to capture meaningful relationships when studying the determinants of an economy's dynamics.

We show that the equilibrium outcome may be extremely sensitive to small differences in long-term growth growth rates, g, and specifically that small differences can have large welfare effects by taking the economy from a Pareto efficient, perfectly competitive, outcome to one in which some industries are competitive and others are not. We study a modified version of our workhorse example from the previous section, given in Table 2. The differences are that there are now 20 firms in each industry, that the asymmetry in industry sizes is not as large as in the previous example, and that there are productivity variations across states also in the large industries. It is straightforward to verify that there is an equilibrium with aggregate consumption

$$C_1 = C_2 = 1.19,$$

and markups

Markups	$s = 1$	$s = 2$
$Q(I_1)$	11	11
$Q(I_2)$	11	4.44
$Q(I_3)$	4.44	11

(33)

and that the efficiency therefore is $e_s = \frac{C_s}{A_s} \equiv \frac{1.19}{1.33} = 0.89$, about 11% below the Pareto efficient outcome in both states.

Type, j	I_j	N	A_1	A_2	α_1	α_2
1	$[0, 0.2)$	20	1	1	0.972	0.972
2	$[0.2, 0.6)$	20	1	2	0.972	1.041
3	$[0.6, 1]$	20	2	1	1.041	0.972
$\bar{A}$					$\bar{A}_1 = 1.33$	$\bar{A}_2 = 1.33$

$$\Phi = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$$

$$\gamma = 6, \qquad \theta = 1.1, \qquad \delta = 0.95,$$

Table 2. Modified economy with three industries and two states.

Now, in an identical economy as the one in Table 2, except for that $\delta = 0.949$ instead of 0.95, only the competitive outcome is an equilibrium, leading to $Q \equiv 1$ and efficient consumption

$$C_1 = C_2 = 1.33.$$

This follows immediately since $N > N^c$ in all industries. Thus, the discontinuity of markups close to N^c, analyzed in Section 3.3, leads to extreme sensitivity of the equilibrium outcome when there is a substantial number of industries in which the number of firms close to N^c. We recall that δ is a function of the long-term growth rate (g), the risk aversion coefficient (γ), and the personal discount rate ($\hat{\delta}$). The equilibrium may therefore be sensitive to these parameters (in other examples, we generate drastic changes in consumption *variation*, not only levels).

Variations in other parameters do typically not have as drastic an effect on the equilibrium. Especially, a property of the economy that turns out to be remarkably insensitive to parameter variations is the number of industries that are competitive. We define

Definition 3. *The* competitiveness, *η, of an equilibrium is defined as the mass of industries that are perfectly competitive,*

$$\eta = \lambda(\{z : Q_s(z) = 1, \forall s\}).$$

We recall from the previous analysis that an industry is either competitive in all states or in no state, so requiring perfect competition in all states is in no way restrictive. We now have

Proposition 6. *The competitiveness of an equilibrium depends on δ and the function N, but not on Φ, α, $\bar{A}$ or θ.*

Thus, somewhat surprisingly, the only real parameter that is important for competitiveness is the long-term growth rate, g. Recall that $\delta = \hat{\delta}(1+g)^{1-\gamma}$. If $\gamma < 1$, a higher long-term growth rate leads to lower competitiveness, i.e., fewer industries in which there is perfect competition. If $\gamma > 1$ on the other hand, a higher long-term growth rate leads to higher competitiveness. Compared with the risk neutral setting ($\gamma = 0$), competitiveness is higher in the economy with risk averse agents as long as the economy is growing in the long term, $g > 0$ (since $\delta(1+g)^{1-\gamma} < \delta(1+g)$). This is not surprising, since future growth is valued less by an agent with a concave utility function, and it is therefore more tempting for firms to deviate from a cooperative outcome when agents have such preferences.

Interestingly, shorter term fluctuations, represented by Φ, *never* matter for the competitiveness. Neither does the distribution of productivity over industries, α, nor aggregate productivity, $\bar{A}$. Upon reflection, the two results — that competitiveness can be extremely sensitive to the long-term growth rate (g) — although it does not depend at all on temporary fluctuations (Φ) may seem puzzling. Indeed, since our set-up is completely general, we could thus have very different competitiveness in two economies that seem to be identical for an arbitrary long (but finite) time period: one economy with a long-term growth rate $g > 0$, and one that grows at rate g for a long time through state transitions in Φ (recall that Φ can be arbitrary large but finite), but that has a long term growth rate of 0. The technical reason why these economies are indeed very different is because of the infinite horizon nature of dynamic collusion games. In long but finite horizon games, collusion is much harder to sustain, as it is in our model when growth only lasts for a finite time period. The importance of long-term growth rates for asset pricing was recently discussed in Parlour et al. (2011). In that paper, long-term growth rates are important because they determine how much investors care about rare disaster events in the far future. The model in this paper provides another mechanism through which long-term growth rates may be crucial in determining the economy's dynamics and asset prices.

5.3 Uniqueness

Our existence result makes no claims with regards to uniqueness. Given an economy, $\mathcal{E}$, there may be multiple equilibria whenever a nonzero measure of firms fails the condition of perfect competition. The parametrization in Table 1 reveals that there is indeed (exactly) one more equilibrium in the economy supported by the equilibrium markups

Markups	$s=1$	$s=2$
$Q(I_1)$	11	2.104
$Q(I_2)$	11	2.605
$Q(I_3)$	11	11

(34)

and leading to aggregate consumption

$$C_1 = 0.974, \qquad C_2 = 0.884.$$

Again, business cycles are endogenous. However, despite the same technology specification, the second equilibrium is very different from the first one. First, although the state with low productivity is the first, aggregate output is the lowest in the second state in this second equilibrium. There is thus a second way to ensure that firms do not deviate from equilibrium strategies, namely to decrease the attractiveness of state 2. We note that the first equilibrium leads to higher output than the second equilibrium in state 2 (1 versus 0.884), whereas the second equilibrium dominates in state 1 (0.974 versus 0.795). So, a combination of the two would be better than either one of them. However, switching between the two is, of course, not feasible. The OBCA in this second equilibrium is

$$OBCA = \frac{\sigma_C}{\sigma_{\bar{A}}} = \frac{0.045}{0.013} = 3.68,$$

lower than the OBCA in the first equilibrium, which suggests that the second equilibrium is better from a welfare perspective. It is indeed easily verified that the second equilibrium Pareto dominates the first in expected utility terms, regardless of the current state.

It turns out that there are multiple equilibria even in the reference economy with no productivity shocks, i.e., in which $A_1 = 1$ also for the first industry. In this economy, one can verify that

Markups	$s=1$	$s=2$
$Q(I_1)$	11	1.634
$Q(I_2)$	11	1.634
$Q(I_3)$	11	11

(35)

with aggregate consumption.

$$C_1 = 1, \qquad C_2 = 0.830,$$

is also an equilibrium. Moreover, an third equilibrium (symmetrically) exists in which markups and consumption are low in state 1. Since productivity (idiosyncratic and aggregate) is constant across states in this case, the business cycle in this equilibrium is not just amplified, it is completely endogenous, and the OBCA is infinite. Thus, truly endogenous business cycles arise because of strategic competition in our model.

To summarize, collusion within an industry produces a unique outcome given the behavior of all other industries and the implied stochastic discount factor (see Proposition 2), but general equilibrium effects may lead to multiplicity of equilibria. This suggests that coordination *across* industries becomes relevant and opens up for a role for policy makers. This is noteworthy because our model shares many properties with standard business cycle models, in which outcomes are efficient. For example, there are no standard frictional costs in our model. Of course, the assumption that the number of firms in each industry is exogenously determined could be viewed as equivalent to assuming a friction in terms of high entry costs. However, even if there were no entry costs, the competitive outcome may not prevail. For example, the outcome with $N = 20$ firms in each industry and aggregate consumption in Table 2 is an equilibrium in the economy with zero costs of entry, since profits would immediately drop to zero if another firm entered an industry, so no firm has an incentive to do so. Moreover, there are no bubbles in the model. Instead, our key deviation from the traditional approach is to assume that intra-industry prices are strategically determined, instead of being determined through a Walrasian mechanism.

6 Asset Pricing

For convenience, we rewrite the previously derived pricing formulas (18,19)

$$V = \Theta\pi,$$

where

$$\Theta = \Lambda_m^{-1}(I - \delta\Phi)^{-1}\Lambda_m - I.$$

Here, $\Lambda_m = diag(m)$, where $m_s = C_s^{-\gamma}$ is the representative agent's marginal utility in state s. Without loss of generality, we assume that $C_1 \leq C_2 \leq \cdots \leq C_S$, so that $m_1 \geq m_2 \geq \cdots \geq m_S$. Thus, low-$s$ states represent recessions where consumption is lower than normal, whereas high-s states represent expansion periods with abnormally high consumption. In the special case of a risk-neutral representative agent, $m = 1$, the pricing formula reduces to

$$V = ((I - \delta\Phi)^{-1} - I)\pi \stackrel{\text{def}}{=} (\Psi - I)\pi,$$

where $\Psi = (I - \delta\Phi)^{-1}$ is the resolvent, which replaces Φ since payoffs are perpetual and discounted. We also define $\hat{\Psi} = \Psi - I = \sum_{i=1}^{\infty} \delta^i \Phi^i$.

A general property of consumption based equilibrium models is that assets that pay off in good states of the world are worth less — and thereby have a higher expected return — than assets that pay off in bad states of the world. We would therefore expect such a result to hold within our setting too.

6.1 Arrow-Debreu Perpetuities

We first study simple Arrow-Debreu "perpetuities" and show that, given some natural technical restrictions on the resolvent, Ψ (and thereby on the transition matrix, Φ), the property above holds within our setting too. Specifically, we study securities 1_j, $1 \leq j \leq S$, that pay off one unit at time $t = 1, 2, \ldots$ if and only if the economy is in state j at time t. The (*ex dividend*) expected return on Arrow-Debreu perpetuity 1_j in state i is denoted μ_{ij}, and it follows from (18,19), that

$$\mu_{ij} = \frac{\left[\Phi\Lambda_m^{-1}\Psi\Lambda_m\right]_{ij}}{[\Lambda_m^{-1}\hat{\Psi}\Lambda_m]_{ij}}. \tag{36}$$

The reason why extra technical restrictions are needed is that in the general case a state that is instantaneously "good" (i.e., has a low m_s) may actually be quite "bad," because it may be very likely that the economy switches to a bad state (a high m_s state) in the next period. Compare an asset that pays off in a good state with one that pays off in a slightly worse state, in such an economy. The first asset, although instantaneously paying off in a better state than the first may actually be viewed as a "bad-state" asset, given the likely short-term dynamics of the economy. This may change the ordering of expected returns of the two assets. To avoid such pathological situations, we need something like a monotone likelihood ratio property (MLRP) to rank the conditional probabilities across states. Following Karlin (1968), we therefore define

Definition 4. *A matrix, B, is said to be* totally positive of order 2 *(TP2) if, for every $i < j$ and $k < \ell$, $B_{ik}B_{jl} - B_{jk}B_{i\ell} \geq 0$.*

For discrete state spaces, total positivity of the transition matrix corresponds to the monotone likelihood ratio property.[15] The following proposition introduces conditions that ensure that the ordering of expected returns of Arrow-Debreu perpetuities is monotone.

Proposition 7. *Define μ_{ij} to be the expected return in state i on the Arrow-Debreu perpetuity 1_j. Assume that the resolvent, Ψ, satisfies the following two conditions:*

$$\Psi \text{ is TP2}, \tag{37}$$

$$\Psi_{i+1,i+1}\Psi_{ii} - \Psi_{i,i+1}\Psi_{i+1,i} \geq \max(\Psi_{i+1,i+1}, \Psi_{i,i}), \qquad 1 \leq i \leq S-1. \tag{38}$$

Then, for all i, for all j and $j' > j$,

$$\mu_{ij'} \geq \mu_{ij},$$

i.e., in each state of the world, the expected return on an Arrow-Debreu perpetuity is higher, the higher state of the world in which it makes its payments.

We note that the condition needed for the Proposition to be satisfied only involves Ψ, not any other parameters of the economy. Also, (38) imposes a stronger condition than total positivity of order 2 on the diagonal of Ψ, since nonnegativity is not sufficient for (38) to be satisfied. In the monopolistic case, $m_s = C_s^{-\gamma} = \bar{A}_s^{-\gamma}$, so aggregate productivity in the form of $\bar{A}$ provides a strong ranking of expected returns: Arrow-Debreu perpetuities that pay off in high-productivity states have higher expected returns than Arrow-Debreu securities that pay off in low-productivity states, and this is true regardless of which state, s, the economy is in.

[15] The close relationship between MLRP and TP2 can be easiest seen from the MLRP of a parametric family of strictly positive, smooth, density functions, $f(x|y)$, being given by $\frac{\partial^2 \log(f(x|y))}{\partial x \partial y} \geq 0$. A Taylor expansion shows that $\frac{\partial^2 \log(f(x|y))}{\partial x \partial y} \approx \frac{1}{\Delta x \Delta y}\frac{1}{f(x|y)}\left(f(x+\Delta x|y+\Delta y)f(x|y) - f(x+\Delta x|y)f(x|y+\Delta y)\right)$ for small Δx and Δy. The TP2 condition is the discrete version of the term within the parentheses, when $f(i\Delta x|j\Delta y)$ is replaced by the matrix elements B_{ij}.

We expect the above intuition to hold for more general assets too. Intuitively, the payouts of any firm in a stationary equilibrium can be viewed as those of a portfolio of Arrow-Debreu perpetuities, and firms that have relatively higher productivity in good states will have portfolios that load up more on high-productivity Arrow-Debreu perpetuities, and therefore also have higher expected returns. We show that this is indeed the case, in Section 6.2. We stress, however, that this intuition crucially depends on the assumption of a monopolistic economy. Without this assumption, the result is invalid, as shown in Section 6.3.

It is easy to show that (37,38) are always satisfied in an economy with two states:

Corollary 1. *In any economy with two states, $S = 2$, (37,38) are satisfied.*

It is possible to get reversal of the results in an economy with three states, however.

Example: Consider the 3-state economy, with

$$\Phi = \begin{bmatrix} 0.5 & 0.1 & 0.4 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{bmatrix},$$

and $\delta = 0.9$, leading to

$$\Psi \approx \begin{bmatrix} 2.9688 & 2.4324 & 4.5988 \\ 1.4063 & 5.1351 & 3.4586 \\ 1.4063 & 2.4324 & 6.1613 \end{bmatrix}.$$

Now, Ψ is not TP2, since, choosing $i = 2, k = 1, j = 3, l = 2$, in Definition 4, we get

$$\Psi_{2,1}\Psi_{3,2} - \Psi_{3,1}\Psi_{2,2} = -3.8007 < 0.$$

Thus, (37) is not satisfied. Given that the marginal utilities are $m_1 = 11$, $m_2 = 10$, $m_3 = 1$, it is straightforward to use (40) to get

$$\mu_{2,1} = 1.99, \qquad \mu_{2,2} = 1.64, \qquad \mu_{2,3} = 2.73,$$

so in state 2, the Arrow-Debreu perpetuity that pays off in state 2, 1_2, has a lower expected returns than the perpetuity that pays off in state 1, 1_1, although state 1 is worse than state 2. This is thus a counterexample to the ranking of expected returns.

It is clear why this nonmonotonicity arises. Although state 2 is slightly better than state 1, once state 1 is reached, there is a decent chance that the economy will quickly move to state 3, since $\Phi_{1,3} = 0.4$, whereas the economy is likely to stay in state 2 for a substantial time period. Therefore, although the marginal utility is higher in state 1 than in state 2, it is a "better" state when dynamics are taken into account.

The example shows that some structure on Φ is indeed needed to obtain an ordering of expected returns. Together with Corollary 1, it also shows that the intuition from the 2-state economy, which has mainly been studied in previous literature, does not in general hold in multi-state economies.

As mentioned, the example above is in some sense pathological since the transition matrix has the property that it may be better to be in a state that short-term is worse. We will therefore focus on economies that satisfy (37,38), going forward. An example of an S-state economy in which the conditions are satisfied is when Φ has the following structure:

$$\Phi = \begin{bmatrix} 1-p & p & \dots & & & & \\ p & 1-2p & p & \dots & & & \\ & p & 1-2p & p & \dots & & \\ & & & \dots & & & \\ & & & & & p & 1-p \end{bmatrix}, \tag{39}$$

i.e., $\Phi_{1,1} = 1-p$, $\Phi_{1,2} = p$, $\Phi_{i,i-1} = p$, $\Phi_{i,i} = 1-2p$, $\Phi_{i,i+1} = p$, $i = 2,\dots,S-1$, $\Phi_{S,S} = 1-p$, $\Phi_{S,S-1} = p$, where $0 < p < 0.5$. In this case, we have

Proposition 8. *In an $S \geq 2$ state economy in which Φ is defined as in (39), conditions (37,38) are satisfied.*

We next study the implications of the previous results for expected returns of general firms. We first focus on the monopolistic case, and show that firms with pro-cyclical demand have higher expected returns that firms with countercyclical demand. We then show that no such ordering occurs in the oligopolistic case.

6.2 Monopolistic Economies

As in Section 3, we study a specific industry, z, but suppress the z dependence for notational convenience. We wish to understand the relationship between how industry product demand and productivity vary with the state of the economy, and the expected returns in that industry. Demand and productivity are conveniently summarized in α. Our intuition is that the analysis carried out for Arrow-Debreu securities should generalize such that firms with relatively higher demand in good states have higher expected returns than firms with relatively higher demand in bad states.

Of course, given the duality between productivity and product demand in our model, in which both are summarized in the function $\alpha(z)$, the previous argument might as well have been given in terms of pro- and countercyclical *productivity*. In line with previous literature, we will use the "product demand" terminology.

The expected return on a general firm that pays $\pi \in \mathbb{R}_+^S$ (where π_s is the payment made in state s), given that the economy is in state i, is denoted by μ_i, and from (18,19), it follows that

$$\mu_i = \frac{\sum_j \left[\Phi \Lambda_m^{-1} \Psi \Lambda_m\right]_{ij} \pi_j}{\sum_j \left[\Lambda_m^{-1} \hat{\Psi} \Lambda_m\right]_{ik} \pi_k}. \tag{40}$$

Given that $\pi_s \propto C_s \alpha_s$, (see (910,20) in the monopolistic economy, we can then write μ as a function of product demand, A, $\mu_i = \mu_i(\alpha) = \mu_i(\alpha_1,\dots,\alpha_S)$.

The following Proposition confirms our intuition that firms with more pro-cyclical normalized demand — in the sense that normalized demand is higher in good states — have higher expected returns that firms with countercyclical normalized demand — in the sense that it is higher in bad states:

Proposition 9. *Consider an industry with demand (productivity) $\alpha \in \mathbb{R}_+^S$ in an economy that satisfies (37,38). The expected return in state i is μ_i. Then, for any state $1 \le i \le S$, there is a cut-off state, $1 \le \bar{s} \le S$, such that*

- $\frac{\partial \mu_i}{\partial \alpha_s} \ge 0$ *for all s such that $s \ge \bar{s}$, and*
- $\frac{\partial \mu_i}{\partial \alpha_s} \le 0$ *for all s such that $s \le \bar{s}$.*

Thus, in each state, i, there is a cutoff state, so that firms with (all else equal) higher product demand in states above the cutoff state have higher expected returns than firms with (all else equal) lower product demand in states above the cutoff state. Similarly, firms with higher product demand in states below the cutoff state have lower expected returns than firms with lower product demand in state below the cutoff state. An immediate corollary is that

Corollary 2. *Increasing product demand in state S (weakly) increases a firm's expected return, whereas increasing product demand in state 1 (weakly) decreases a firm's expected return.*

These results verify our intuition about the relationship between demand cyclicality and expected returns.

6.3 Oligopolistic Economies

In this section we show that, since monopolistic markups cannot be sustained in all states in the oligopolistic setting, the link between product demand cyclicality and expected returns breaks down in the oligopolistic setting.[16] The effect is best shown in an example. For simplicity, consider a 25-state economy with a monopolistic equilibrium (such that $\bar{A} \equiv C$), and two zero measure industries in which there may be strategic competition. The first industry has procyclical demand in that α increases with aggregate consumption, whereas the second industry is countercyclical in that α decreases with aggregate consumption. The parameters in the economy are given by:

$$C_1 = 1.0317, \qquad C_s = 0.5 + 0.4s,\ 2 \le s \le S, \tag{41}$$

$\theta = 1.1, \gamma = 3$ and $\delta = 0.9$. In the oligopolistic case we assume that $N = 9$ in both industries, whereas in the monopolistic case, $N = 1$. The productivity functions of the two firms, α^1 and α^2, respectively, are given in the appendix. Finally, as before, the transition matrix is given by

$$\Phi = \begin{bmatrix} 1-p & p & \dots & & & & \\ p & 1-2p & p & \dots & & & \\ & p & 1-2p & p & \dots & & \\ & & & \dots & & & \\ & & & & & p & 1-p \end{bmatrix}, \tag{42}$$

where $p = 0.1$, i.e., $\Phi_{1,1} = 1-p$, $\Phi_{1,2} = p$, $\Phi_{i,i-1} = p$, $\Phi_{i,i} = 1-2p$, $\Phi_{i,i+1} = p$, $i = 2, \dots, S-1$, $\Phi_{S,S} = 1-p$, $\Phi_{S,S-1} = p$. The parameters are chosen to make the results easy to interpret in a figure.

[16] As we have previously mentioned, our technology shocks can also be interpreted as demand shocks.

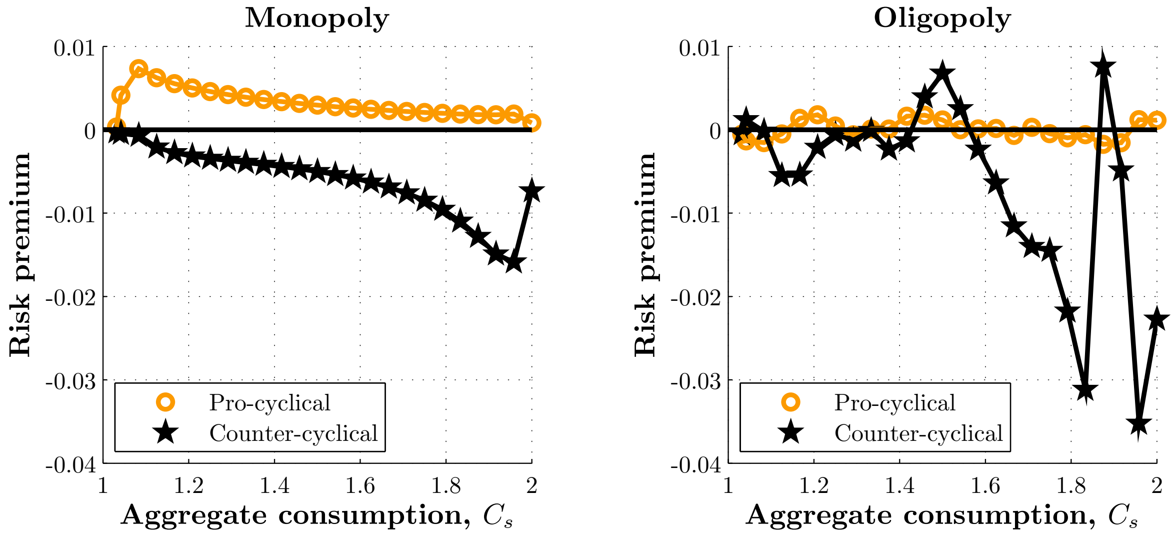

Figure 4. This graph plots the state-contingent risk premia for pro-and counter-cyclical industries under monopolistic behavior (left panel) and oligopolistic behavior (right panel).

In the left panel of Figure 4, the risk premium, i.e., the expected return above the risk-free rate, is shown for the two firms when they are monopolistic. Along the lines of the results in the previous section, the procyclical firm has a positive risk premium in all states of the world, whereas the counter cyclical firm has a negative risk premium. This is just as expected, since the procyclical firm tends to generate dividends in good states and therefore carries a higher risk-premium than the market in all states, whereas the opposite is true for the firm with countercyclical product demand.

This intuition breaks down for the firms in the oligopolistic case, however. As seen in the right panel of Figure 4, there is no simple relationship between risk premia of the procyclical and countercyclical firms over the business cycle in the oligopolistic case. In some states the countercyclical firm has higher expected returns than the procyclical firm, and in other states it is lower. Both firms have positive as well as negative risk premia, depending on the state.

This complex dependence of risk premia over the business cycle arises already in the case with a few strategic industries in an otherwise monopolistic economy. In the general case, when the market price of risk is also affected by oligopolistic industries, the relationship between returns and procyclicality will be even more complex. In this case, the strategic competition will also be influenced by the aggregate market power in different states, q_s, which introduces an additional wedge between product demand cyclicality and returns in the right panel of Figure 4.

The main asset pricing implication of this analysis is that given the complex dynamics of expected returns across the business cycle that may arise when firms compete strategically, it should be fruitful to use industry characteristics, as well as further firm characteristics, as conditioning variables when explaining expected returns. For example, the number of firms in an industry — or other measures of industry concentration — should provide useful information for the cross section of returns. Also, the markup a firm charges over the business cycle, e.g., measured by some type of profitability proxy, as well as proxies of product demand cyclicality, should be valuable conditioning variables.

To make this point explicit, we simulate 1,000 firms in different industries in our model to study how well firm and industry characteristics explain the cross section of expected returns. The example is necessarily very stylized, but captures the intuition of our paper. We focus on unconditional expected returns and use a similar specification as the one just studied. Again, we consider a 25-state economy with a monopolistic equilibrium (such that $\bar{A} \equiv C$), and and focus on zero measure industries in which there may be strategic competition. The parameters in the economy are given by:

$$C_s = 1 + 0.4s, \; 1 \leq s \leq S, \tag{43}$$

$\theta = 1.1, \gamma = 8$ and $\delta = 0.9$. Each firm, i, has productivity $\alpha_s = 1 + 9\zeta_{is}$, in state s, where ζ_{is} are uniformly $(0,1)$ distributed and are independent across firms and states. Further, the number of firms in the industry in which firm i operates, $N_i = 6 + 4\zeta_i^N$, where ζ_i^N is uniformly $(0,1)$ distributed, and independent across firms. Further N's and α's are jointly independent. The transition matrix between states is again tridiagonal, but we vary the probabilities across states, such that $\Phi_{i,i+1} = p_i^1$, $\Phi_{i,i-1} = p_i^2$, $\Phi_{i,i} = 1 - p_i^1 - p_i^2$. Thus, the structure of Φ is as before in that transitions only occur to neighboring states, but we allow the transition probabilities to vary with the state. The reason why we use a slightly more complex transition matrix is that the effects we wish to show are somewhat mitigated — although still present — in the special case of constant probabilities.[17] The coefficients p_i^1 and p_i^2 are given in the appendix.

We expect the number of firms in an industry, N_i, to be important for expected returns, because it will determine the competitiveness in that industry, and ultimately in which states of the world profits are generated. We also expect additional information to be captured in the productivity and profits across states, α_{is} and π_{is}, respectively, and especially by the relationship between these two variables. One approach, which we will take, is to scale profits with productivity, and study the variable $\omega_{is} = \frac{\pi_{is}}{\alpha_{is}}$ across states for a given firm, i. We note that ω_{is} should in principle be straightforward to estimate in practice, since π_{is} is proxied by various profit measures on a firm's income sheet (e.g., by net income or NOPLAT), and α_{is} could be captured by estimating the product demand function in an industry. We have experimented with summary statistics of ω_i across states. It turns out that the mean of ω_i is highly correlated with N_i, since they both capture average competitiveness across states. Similarly, the volatility of ω_i across states also turns out to be highly correlated with N_i and thereby capturing similar information. The skewness of ω, however, is less correlated and — as it turns out — also captures significant information about the cross section of returns.[18] We therefore define $\Omega_i = skew(\omega_{is})$.

We regress unconditional expected returns on unconditional market betas, β_i, number of firms in the industry, N_i, and skewness of profits over productivity, Ω_i, across firms,

$$\mu_i = a_i^1\beta_i + a_i^2 N_i + a_i^3\Omega_i + \epsilon_i, \qquad i = 1, \ldots, 1,000,$$

using an ordinary least square regression. We also carry out univariate regressions for each of the variables. The results are shown in Table 3. We see that in the univariate regressions, both N_i and Ω_i are superior in determining expected returns, with r-squares of about 0.6, compared with market beta which only has an r-square of 0.12. In the multivariate regressions,

[17] Technically, constant probabilities leads to the one case in which the CAPM works quite well, because they imply a uniform stationary distribution.

[18] It is outside of the scope of this study to investigate why this is the case.

Univariate				
Coefficient		a_1	a_2	a_3
		0.0056***	0.0034***	0.012***
R-square		0.12	0.59	0.60
Multivariate				
Coefficient		a_1	a_2	a_3
		-0.0017**	0.023***	0.024***
R-square	0.75			
Correlation	a_1	1	0.36	0.53
	a_2	0.36	1	0.62
	a_3	0.53	0.62	1

Table 3. Cross sectional regression of unconditional expected returns on market betas, number of firms, and the skewness of profits divided by productivity.

the coefficient on market beta is even negative. The total r-square in this regression reaches 0.75, and is only marginally lower when market beta is left out (about 0.73, not reported in the table). The correlation between beta and number of firms is quite low, about 0.36, whereas the other correlations are somewhat higher, between 0.5-0.6. Thus, in line with the intuition of our study, firm and industry characteristics together provide a good characterization of the cross section of expected returns, capturing non-redundant but somewhat overlapping pieces of information.

7 Concluding Remarks

We have developed general equilibrium in a dynamic economy with a continuum of industries each of which comprises a finite number of firms. The framework is quite tractable, and the strategic interaction between firms in each industry is straightforward to characterize. We establish the existence of general equilibrium and establish dynamic properties of the economy including equilibrium markups and attendant asset prices.

The central premise of our model is that firms, maximizing shareholder value, are not always price takers but can be price setters. High prices in an industry can be sustained if firms value the future flow of profits over any immediate increases in market share garnered by undercutting. Of course, the rate at which future profits are discounted depends both on the representative agent's risk aversion and on the behavior of the aggregate economy.

The strategic behavior of firms leads to two distinct type of predictions. First, we develop various general equilibrium effects that can interpreted in light of the macro-economy. Even in an economy with no aggregate uncertainty, if the relative productivity of various industries changes, so does their ability to sustain collusive outcomes. These changes can affect both the level and the volatility of aggregate consumption; in short our model exhibits endogenous volatility. Second, we show how industry characteristics are related to asset prices. Specifically, industry characteristics (such as concentration and productivity) affect firms' ability to set prices strategically. An immediate implication of this is that these inter-industry differences should help to explain asset prices.

While conceptually simple, the model is sufficiently rich that it presents many avenues for future research, both on the strategic behavior of firms and on the general equilibrium properties. First, we take the number of firms in each industry as exogenous. However, a natural extension would be to model strategic entry and exit. We speculate that this would lead even more strategic volatility. Second, while we present various stylized examples, it would be natural to structurally calibrate or estimate the model. For example, it is worth noting that the consumption aggregator is highly non-linear. This means that a linear aggregator (such as a price weighted one) will consistently underestimate the volatility of the utility an agents enjoys from a consumption stream. It is an empirical question as to whether this framework provides new empirical insight into the relationship between the volatility of asset prices and aggregate consumption, however it is an interesting avenue for further research. Finally, our model has potentially new implications for policy actions. Indeed, the existence of multiple equilibria suggests that government intervention may be beneficial.

A Optimal Expenditure Share

Maximizing the objective function in state s given by equation 3 subject to the budget constraint results in the following Lagrangian:

$$L_s = \left(\int_0^1 c_s(z)^{\frac{\theta-1}{\theta}} dz\right)^{\frac{\theta}{\theta-1}} - \lambda\left(y_s - \int_0^1 p_s(z)c_s(z)dz\right). \tag{44}$$

The first-order conditions with regards to consumption for some arbitrary $c_s(z')$ and $c_s(z'')$ imply:

$$c_s(z')^{\frac{\theta-1}{\theta}-1}\left(\int_0^1 c_s(z)^{\frac{\theta-1}{\theta}} dz\right)^{\frac{\theta}{\theta-1}-1} = \lambda p_s(z'), \tag{45}$$

$$c_s(z'')^{\frac{\theta-1}{\theta}-1}\left(\int_0^1 c_s(z)^{\frac{\theta-1}{\theta}} dz\right)^{\frac{\theta}{\theta-1}-1} = \lambda p_s(z''). \tag{46}$$

Using these two equations results in:

$$\frac{\frac{1}{\theta}c_s(z')^{\frac{-1}{\theta}}}{\frac{1}{\theta}c_s(z'')^{\frac{-1}{\theta}}} = \frac{p_s(z')}{p_s(z'')}. \tag{47}$$

Since z' was arbitrary, this gives us:

$$c_s(z) = c_s(z'')\frac{p_s(z'')^\theta}{p_s(z)^\theta}. \tag{48}$$

Plugging this expression $c_s(z)$ into the budget equation yields:

$$y_s = \int_0^1 p_s(z)c_s(z'')\frac{p_s(z'')^\theta}{p_s(z)^\theta}dz. \tag{49}$$

Solving for $c_s(z'')$ gives us:

$$c_s(z'') = \frac{1}{p_s(z'')^\theta}\frac{y_s}{P_s^{1-\theta}}. \tag{50}$$

where

$$P_s = \left(\int_0^1 p_s(z)^{1-\theta}dz\right)^{\frac{1}{1-\theta}} \tag{51}$$

represents the appropriate price index for *CES* preferences.

B Proofs

Proof of Proposition 1

Since the wage rate is a free variable, we can always normalize it in such a way that $P_s = 1$. Taking $Q_s(z)$ as given: the wage rate is implicitly determined by

$$\left(\int_0^1 \left[w_s l_s(z) Q_s(z)\right]^{1-\theta} dz\right)^{\frac{1}{1-\theta}} = 1. \tag{52}$$

Solving for w_s yields:

$$w_s = \frac{1}{\left(\int_0^1 l_s(z)^{1-\theta} Q_s(z)^{1-\theta} dz\right)^{\frac{1}{1-\theta}}}. \tag{53}$$

Using the definition of $\alpha_s(z) = \frac{l_s(z)^{1-\theta}}{\bar{A}_s^{\theta-1}}$ (see equation 9) and the power mean definition we can write this as:

$$w_s = \frac{\bar{A}_s}{M_{1-\theta}(Q_s)}. \tag{54}$$

Plugging the demand function, $c_s(z)$ (see equation 4), into the profit equation (see equation 6), yields profits of:

$$\pi_s(z) = [p_s(z) - w_s l_s(z)] \frac{1}{p_s(z)^\theta} y_s. \tag{55}$$

Writing $p_s(z) = w_s l_s(z) Q_s(z)$ and using the definition of $\alpha_s(z) = \frac{l_s(z)^{1-\theta}}{\bar{A}_s^{\theta-1}}$ yields:

$$\pi_s(z) = w_s^{1-\theta} l_s(z)^{1-\theta} [Q_s(z) - 1] \frac{1}{Q_s(z)^\theta} y_s \tag{56}$$

$$= \alpha_s(z) \frac{Q_s(z) - 1}{Q_s(z)^\theta} \bar{A}_s^{\theta-1} y_s w_s^{1-\theta} \tag{57}$$

$$= \alpha_s(z) \frac{Q_s(z) - 1}{Q_s(z)^\theta} \frac{y_s}{[M_{1-\theta}(Q_s)]^{1-\theta}}, \tag{58}$$

where the last line uses the definition of $w_s = \frac{\bar{A}_s}{M_{1-\theta}(Q_s)}$. Total income satisfies:

$$y_s = w_s + \int \pi_s(z) dz \tag{59}$$

$$= w_s + y_s \int \alpha_s(z) \frac{Q_s(z) - 1}{Q_s(z)^\theta} \frac{1}{[M_{1-\theta}(Q_s)]^{1-\theta}} dz. \tag{60}$$

Solving for y_s yields (using w_s):

$$y_s = \frac{\bar{A}_s}{M_{1-\theta}(Q_s)} \frac{1}{\left(1 - \frac{1}{[M_{1-\theta}(Q_s)]^{1-\theta}} \int \alpha_s(z) \frac{Q_s(z)-1}{Q_s(z)^\theta} dz\right)} \tag{61}$$

$$= \frac{\bar{A}_s}{M_{1-\theta}(Q_s)} \frac{1}{\left(1 - \frac{1}{[M_{1-\theta}(Q_s)]^{1-\theta}} \left[\int \alpha_s(z) Q_s(z)^{1-\theta} dz - \int \alpha_s(z) Q_s(z)^{-\theta} dz\right]\right)}. \tag{62}$$

Employing the definition of power means results in:

$$y_s = \frac{\bar{A}_s}{M_{1-\theta}(Q_s)} \frac{1}{\left(1 - \frac{1}{[M_{1-\theta}(Q_s)]^{1-\theta}} \left[[M_{1-\theta}(Q_s)]^{1-\theta} - [M_{-\theta}(Q_s)]^{-\theta}\right]\right)} \tag{63}$$

$$= \frac{\bar{A}_s}{M_{1-\theta}(Q_s)} \frac{1}{\left(1 - \frac{[M_{1-\theta}(Q_s)]^{1-\theta}}{[M_{1-\theta}(Q_s)]^{1-\theta}} + \frac{[M_{-\theta}(Q_s)]^{-\theta}}{[M_{1-\theta}(Q_s)]^{1-\theta}}\right)} \tag{64}$$

$$= \frac{\bar{A}_s}{M_{1-\theta}(Q_s)} \frac{[M_{-\theta}(Q_s)]^\theta}{[M_{1-\theta}(Q_s)]^{\theta-1}} = \bar{A}_s \left[\frac{M_{-\theta}(Q_s)}{M_{1-\theta}(Q_s)}\right]^\theta. \tag{65}$$

Proof of Lemma 1

The lemma is a special case of the following general lemma (by choosing $b = \Theta^T \iota_j$).

Lemma 6. *Consider a strictly positive vector $\pi^m \in \mathbb{R}^S_{++}$, a strictly positive matrix $\Theta \in \mathbb{R}^{S\times S}_{++}$, and a scalar $n \in \mathbb{R}_{++}$. Then there is a unique $\xi \in \mathbb{R}^S_+$ so that for all strictly positive $b \in \mathbb{R}^S_{++}$,*

$$\begin{aligned}
&\xi = \arg\max_x \; b^T x, \;\; s.t.,\\
&x \leq \pi^m,\\
&0 \leq (\Theta - nI)\, x.
\end{aligned}$$

For each s, the solution has either the first or the second constraint binding, i.e., for each s, $\xi_s = \pi^m_s$ or $n\xi_s = \Theta\xi_s$.

Proof: Let $x < y$ denote that $x \leq y$ and $x \neq y$. Also, define $z = x \vee y \in \mathbb{R}^S$, where $z_s = \max(x_s, y_s)$ for all s. Clearly, $x \leq x \vee y$, where the inequality is strict if there is an s such that $y_s > x_s$. Finally, define the set $K = \{x : 0 \leq x,\, x \leq \pi^*, nx \leq \Theta x\}$. Note that K is compact.

Now, there is a unique maximal element of K, that is, there is a unique $\xi \in K$, such that for all $x \in K$ such that $x \neq \xi$, $\xi > x$. This follows by contradiction, because assume that there are two distinct maximal elements, y and x, then clearly $z = x \vee y$ is strictly larger that both x and y. Now, it is straightforward to show that $z \in K$. The only condition that is not immediate is that $\Theta z \geq Nz$. However, this follows from $\Theta(x \vee y) \geq \Theta x \vee \Theta y \geq nx \vee ny = n(x \vee y) = nz$.

Now, since b is strictly positive, it is clear that ξ is indeed the unique solution to the optimization problem regardless of b. That one of the constraint is binding for each s also follows directly, because assume to the contrary that neither constraint is binding in some state s. Then ξ_s can be increased without violating either constraint in state s and, moreover, the constraints in all the other states will actually be relaxed, so such an increase is feasible. Further, since $b_s > 0$, it will also increase the objective function, contradicting the assumption that ξ is optimal.

Proof of Proposition 2

Follows immediately from Lemma 1.

Proof of Lemma 2

By definition: $V = \Lambda_\pi(\Gamma - \mathbf{1})$, so from (20), $\Lambda_\pi(\Gamma - \mathbf{1}) = \delta\Lambda_m^{-1}\Phi\Lambda_m(\pi + \Lambda_\pi(\Gamma - \mathbf{1}))$, leading to $\Gamma - \mathbf{1} = \delta\Lambda_\pi^{-1}\Lambda_m^{-1}\Phi\Lambda_m\Lambda_\pi\Gamma$. Now, observing (from (20)) that $\Lambda_\kappa = \Lambda_\pi\Lambda_m$, the result follows immediately.

Proof of Proposition 3

Let $n = N - 1$ and $K^*(n) \stackrel{\text{def}}{=} \{x : 0 \leq x, nx \leq \Theta x\}$. Now, $nx \leq (\Lambda_m^{-1}(I - \delta\Phi)^{-1}\Lambda_m - I)$ is equivalent to $Ny \leq (I - \delta\Phi)^{-1}y$, where $y = \Lambda_m x \in \mathbb{R}^S_+$. We first show that $K^*(n) = \{0\}$ when $N > \frac{1}{1-\delta}$, which immediately implies that the only solution to the optimization problem in Lemma 1 is indeed the competitive outcome. Define the matrix norm $\|A\| = \sup_{x\in\mathbb{R}^S\setminus\{0\}} \frac{\|Ax\|}{\|x\|}$, where the l^1 vector norm $\|y\| = \sum_s |y_s|$ is used. Since Φ is a stochastic matrix, $\|\Phi^i\| = 1$ for all i and using standard norm inequalities it therefore follows immediately that

$$\|(I - \delta\Phi)^{-1}\| = \left\|\sum_0^\infty \delta^i\Phi^i\right\| \leq \sum_0^\infty \delta^i\|\Phi^i\| = \frac{1}{1-\delta},$$

and thus $\|(I-\delta\Phi)^{-1}y\| \le \frac{1}{1-\delta}\|y\|$. Now, $Ny \le (I-\delta\Phi)^{-1}y$ implies that $N\|y\| \le \|(I-\delta\Phi)^{-1}y\|$, and therefore it must be the case that $N \le \frac{1}{1-\delta}$, for the inequality to be satisfied for a non-zero y. Now, consider the case when $N = \frac{1}{1-\delta}$. Since $y = 1$ is an eigenvector to Φ with unit eigenvalue, it is also an eigenvector to $(I-\delta\Phi)^{-1}$ with corresponding eigenvector $\frac{1}{1-\delta}$, leading to $x = \Lambda_m^{-1}1 = m^{-1}$. It is easy to show that this is the unique (up to multiplication) nonzero solution. Given the properties of Φ, the Perron-Frobenius theorem implies that this is indeed the *only* eigenvector with unit eigenvalue, and therefore also the only eigenvector to $(I-\delta\Phi)^{-1}$ with eigenvalue $\frac{1}{1-\delta}$. Now, take an arbitrary $y \in \mathbb{R}_+^S \setminus \{0\}$ as a candidate vector to satisfy the inequality, i.e., such that $z = (I-\delta\Phi)^{-1}y$ satisfies $z_i \ge Ny_i = \frac{1}{1-\delta}y_i$ for all i. Then, since $\|(I-\delta\Phi)^{-1}\| = \frac{1}{1-\delta}$, it follows that $\sum_i z_i \le \frac{1}{1-\delta}\sum_i y_i$. The two inequalities can only be satisfied jointly if $z_i = \frac{1}{1-\delta}y_i$ for all i, and thus y is the already identified eigenvector. Thus, $K^*\left(\frac{1}{1-\delta}\right) = \{\iota m^{-1}, \iota \ge 0\}$. It follows immediately from the definition of the λ vector that the maximal ι that satisfies $\iota m_s^{-1} \le \pi_s^* = q_s C_s \alpha_s$ for all s is $\min_s \lambda_s$, leading to the given form of the profit vector.

Proof of Lemma 3

If $\kappa_s = k$, the diagonal matrix Λ_κ becomes $\Lambda_\kappa = kI$ so that we obtain for Γ (see (29)):

$$\Gamma = (I-\delta\Phi)^{-1}\mathbf{1} = \frac{1}{1-\delta}\mathbf{1} = N^c\mathbf{1}. \tag{66}$$

This is because the eigenvalue of $(I-\delta\Phi)^{-1}$ associated with the eigenvector of $\mathbf{1}$ is given by $\frac{1}{1-\delta}$ (see Proof of Proposition 3). So, $N^m = \min_s(\Gamma_s) = N^c$.

Proof of Proposition 4

(1,2) follow from the definition of K in the proof of Lemma 1. It immediately follows that the set K is decreasing in N and increasing in each of α_s, which in turn immediately implies (1,2).

(3) follows from (1), and the fact that $\pi_s > 0$ for all s when the number of firms is N^c.

(4) follows from (1) and that $\pi_s = m_s^{-1}\pi_s^m m_s$ for the s that minimizes μ_s (see Proposition 3).

(5) follows from the fact that the objective function in Lemma 1 is a continuous function of all parameters and that (as long as N is strictly below N^c) the set K is compact, and depends continuously on all parameters, in the sense that if K and K' are defined for two sets of parameter values, then $D(K,K')$ approaches zero when the parameter values that define K' approach those that define K. Here, $D(K,K') = \sup_{x\in K'}\inf_{y\in K}|x-y|$.

Proof of Proposition 5

We wish to prove the proposition with a fixed point argument, and therefore define a fixed point relationship for the markup function, Q, which ensures that it defines an equilibrium. We define $R \stackrel{\text{def}}{=} \bar{N}\times[c,C]^S$, where $\bar{N} = \{1,2,\lfloor N_c\rfloor+1\}$, with elements $x = (n,\alpha_1,\ldots,\alpha_S) \in R$. We will then work with functions $Q^0 : R \to [0,1]^S$, and given such a function, the transformation to the standard markup function is given by $Q_s(z) = Q_s^0(\min(N(z),\lfloor N_c\rfloor+1),\alpha_1(z),\ldots,\alpha_S(z))$. The reason why we work with the canonical domain, R, rather than $S\times[0,1]$, is that compactness properties needed for a fixed point argument are easier obtained in this domain. Given a function, $Q^0 : R \to \left[1, \frac{\theta}{\theta-1}\right]^S$, we define

$$p_s^0 = M_{-\theta}\left(Q_s\right) = \left(\int \alpha_s(z) Q_s\left(z\right)^{-\theta} dz\right)^{\frac{1}{-\theta}} = \left(\int_{x\in R} x_{s+1} Q^0(x)^{-\theta} dF(x)\right)^{\frac{1}{-\theta}}, \tag{67}$$

$$p_s^1 = M_{1-\theta}\left(Q_s\right) = \left(\int \alpha_s(z) Q_s\left(z\right)^{1-\theta} dz\right)^{\frac{1}{1-\theta}} = \left(\int_{x\in R} x_{s+1} Q^0(x)^{1-\theta} dF(x)\right)^{\frac{1}{1-\theta}}. \tag{68}$$

It follows immediately that the mapping from Q^0 to p_0 and p_1 is continuous (in L^1 topology) and since $\int \alpha(z)dz = 1$, that p_s^0 and p_s^1 lie in $[1, \theta/(\theta-1)]$. From (13), it follows that

$$C_s = \bar{A}_s \left(\frac{p_1}{p_0}\right)^{\theta}, \tag{69}$$

and from (20) that

$$\pi_s^m = \frac{1}{p_1^{1-\theta}} \frac{(\theta-1)^{\theta-1}}{\theta^\theta} \alpha_s C_s = \frac{1}{p_1^{1-\theta}} \frac{(\theta-1)^{\theta-1}}{\theta^\theta} x_{s+1} C_s. \tag{70}$$

Now, for each z, given $\pi^m \in \mathbb{R}_+^S$, the program in Lemma 1 provides a continuous mapping from π^m to

$$\pi_s \in \prod_1^S [0, \pi_s^m]. \tag{71}$$

We use (15) to define the operator $\mathcal{F}$, which operates on functions, and which is given by:

$$Q_s^1(x) = (\mathcal{F}(Q^0)(x))_s = 1 + \frac{p_1(s)^{1-\theta}}{C_s x_{s+1}} (Q_s^0(x))^\theta \pi_s.$$

Since each operation in (67-71) is continuous, it follows that $\mathcal{F}$ is a continuous operator (in $L^1(\mathbb{R}^{1+S})$-norm). Further, it also follows that if $Q_s^0(x) \in \left[1, \frac{\theta}{\theta-1}\right]$, then since $0 \leq \pi \leq \pi^m$, $1 \leq Q_s^1(x) \leq 1 + \frac{(\theta-1)^{\theta-1}}{\theta^\theta}(Q_s^0)^\theta \leq \frac{\theta}{\theta-1}$. Define, Z as the set of all functions, $Q : R \to [1, \theta/(\theta-1)]^S$, such that Q is nonincreasing in its first argument and nondecreasing in all other arguments. Then, from what we have just shown, together with Proposition 4, it follows that $\mathcal{F}$ is a continuous operator that maps Z into itself. We also have

Lemma 7. *Z is convex and compact.*

We prove that the set, W, of nondecreasing functions $f : [0,1] \to [0,1]$, is convex and compact. The generalization to functions with arbitrary rectangular domains and ranges, $f : \prod_1^N [a_i, b_i] \to \prod_1^M [c_i, d_i]$, is straightforward, as is the generalization to functions that are nonincreasing in some coordinates and nondecreasing on others (as is Z). Convexity is immediate. For compactness, we show that every sequence of functions $f^n \in W$, $n = 1, 2, \ldots$, has a subsequence that converges to an element in W. First, note that W is closed, since a converging (Cauchy) sequence of nondecreasing functions necessarily converges to a nondecreasing function. To show compactness, define the corresponding sequence of vectors $g^n \in [0,1]^{2^j}$, for some $j \geq 1$, by $g_k^n = f_n(2^{-j}k)$, $k = 0, 1, \ldots 2^j - 1$. Now, since $[0,1]^{2^j}$ is compact it follows that there is a subsequence of $\{f^n\}$, $\{f^{n_m}\}$ that converges at each point $2^{-j}k$, to some $g^* \in [0,1]^{2^j}$. Define the function $h^j : [0,1] \to [0,1]$ by $h^j(x) = g_k^*$, for $2^{-j}k \leq x < 2^{-j}(k+1)$, which is obviously also in W. Next, take the sequence $\{f^{n_m}\}$, and use the same argument to find a subsequence that converges in each point $2^{-(j+1)}k$, $k = 0, \ldots, 2^{j+1} - 1$, and the corresponding function $h^{j+1}(x)$. By repeating this step, we obtain a sequence of functions in W, $h^j, h^{j+1}, \ldots$, such that for $m > j$,

$$\int_0^1 |h^m(x) - h^j(x)| dx \leq \sum_k (g_{k+1}^j - g_k^j) 2^{-j} \leq 2^{-j}.$$

Thus, $h^j, h^{j+1}, \ldots$ forms a Cauchy-sequence, which consequently converges to some function $h^* \in W$. Take a subsequence of the original sequence of functions, $\{f^{n_j}\}$, such that $\int |f^{n_j} - h^j| dx \leq 2^{-j}$. Then, for $m > j$, since

$$\begin{aligned}\int_0^1 |f^{n_m}(x) - f^{n_j}(x)| dx &= \int_0^1 |f^{n_m}(x) + h^m(x) - h^m(x) + h^j(x) - h^j(x) - f^{n_j}(x)| dx \\ &\leq \int_0^1 |f^{n_m}(x) - h^m(x)| dx + \int_0^1 |f^{n_j}(x) - h^j(x)| dx \\ &+ \int_0^1 |h^m(x) - h^j(x)| dx \\ &\leq 3 \times 2^{-j},\end{aligned}$$

$\{f^{n_j}\}$ is also a Cauchy sequence and converges to $h^* \in W$. Thus, W is compact and the lemma is proved. Given Lemma 7 and the continuity of $\mathcal{F}$, a direct application of Schauder's fixed point theorem

implies that there is a $Q^* \in Z$, such that $\mathcal{F}(Q^*) = Q^*$. Now, given such a Q^*, and its associated π^m defined by (70), and given the functions, $N(z)$ and $\alpha_s(z)$, $0 \leq z \leq 1$, Lemma 1 can be used to construct $Q_s(z)$. Since Q and Q^* have the same distributional properties, and C, p_0 and p_1, only depend on distributional properties, it immediately follows that Q constitutes an equilibrium. We are done.

Proof of Proposition 6

Follows immediately from Proposition 3.

Proof of Proposition 7

From (36), it follows that

$$\mu_{ij} = \frac{\left[\Phi \Lambda_m^{-1} \Psi \Lambda_m\right]_{ij}}{[\Lambda_m^{-1}(\Psi - I)\Lambda_m]_{ij}} = \frac{\left[\Phi + \Lambda_m^{-1} \hat{\Psi} \Lambda_m\right]_{ij}}{[\Lambda_m^{-1} \hat{\Psi} \Lambda_m]_{ij}} = \frac{\Phi_{ij} + \sum_k \Phi_{ik} \frac{1}{m_k} \hat{\Psi}_{kj} m_j}{\frac{1}{m_i} \hat{\Psi}_{ij} m_j}, \tag{72}$$

where $\hat{\Psi} = \Psi - I$. Now, since $\hat{\Psi} = \Psi - I = \delta\Phi + \delta^2\Phi^2 + \cdots = \delta(\Phi + \Phi(\delta\Phi + \cdots)) = \delta(\Phi + \Phi\hat{\Psi})$ it follows that

$$\frac{\Phi_{ij} + \sum_k \Phi_{ik} \hat{\Psi}_{kj}}{\hat{\Psi}_{ij}} = \frac{1}{\delta},$$

which we use to rewrite

$$\begin{aligned}\mu_{ij} &= m_i \left(\frac{\Phi_{ij} \frac{1}{m_i} + \sum_k \Phi_{ik} \frac{1}{m_i} \hat{\Psi}_{kj}}{\hat{\Psi}_{ij}} + \frac{\Phi_{ij} \left(\frac{1}{m_j} - \frac{1}{m_i}\right) + \sum_k \Phi_{ik} \left(\frac{1}{m_k} - \frac{1}{m_i}\right) \hat{\Psi}_{kj}}{\hat{\Psi}_{ij}} \right) \\ &= \frac{1}{\delta} + m_i \sum_k r_i(k) \Phi_{ik} \frac{\Psi_{kj}}{\hat{\Psi}_{ij}},\end{aligned}$$

where $r_i(k) \stackrel{\text{def}}{=} \frac{1}{m_k} - \frac{1}{m_i}$ because of the ordering of the states is increasing in k for each i, such that $r_i(k) \leq 0$ if $k < i$, $r_i(k) \geq 0$ if $k > i$ and $r_i(i) = 0$. We now show that μ_{ij} is (weakly) increasing in j for each i, i.e., that the procyclicality ordering is indeed valid for the Arrow-Debreu securities. To see this, note that

$$\mu_{i,j+1} - \mu_{i,j} \geq 0 \Leftrightarrow \sum_k r_i(k)\Phi_{ik}\left(\frac{\Psi_{k,j+1}}{\hat{\Psi}_{i,j+1}} - \frac{\Psi_{kj}}{\hat{\Psi}_{ij}}\right) \geq 0$$

$$\Leftrightarrow \sum_{k<i} r_i(k)\Phi_{ik}\left(\frac{\Psi_{k,j+1}}{\hat{\Psi}_{i,j+1}} - \frac{\Psi_{kj}}{\hat{\Psi}_{ij}}\right) + \sum_{k>i} r_i(k)\Phi_{ik}\left(\frac{\Psi_{k,j+1}}{\hat{\Psi}_{i,j+1}} - \frac{\Psi_{kj}}{\hat{\Psi}_{ij}}\right) \geq 0.$$

Now, for $i < j$, $\hat{\Psi}_{ij} = \Psi_{ij}$, so the condition can be written

$$\sum_{k<i} r_i(k)\Phi_{ik}\left(\frac{\Psi_{k,j+1}}{\Psi_{i,j+1}} - \frac{\Psi_{kj}}{\Psi_{ij}}\right) + \sum_{k>i} r_i(k)\Phi_{ik}\left(\frac{\Psi_{k,j+1}}{\Psi_{i,j+1}} - \frac{\Psi_{kj}}{\Psi_{ij}}\right) \geq 0.$$

Now, since Ψ is TP2, $\frac{\Psi_{k,j+1}}{\Psi_{i,j+1}} - \frac{\Psi_{kj}}{\Psi_{ij}}$ is nonpositive for $k < i$ and nonnegative for $k > i$, so since $r_i(k)$ is nonpositive for the first term and nonnegative for the second term, the total expression is indeed weakly positive. An identical argument can be made for $i > j+1$. For $i = j$, the expression becomes

$$\sum_{k<i} r_i(k)\Phi_{ik}\left(\frac{\Psi_{k,i+1}}{\Psi_{i,i+1}} - \frac{\Psi_{ki}}{\Psi_{ii}-1}\right) + \sum_{k>i} r_i(k)\Phi_{ik}\left(\frac{\Psi_{k,i+1}}{\Psi_{i,i+1}} - \frac{\Psi_{ki}}{\Psi_{ii}-1}\right) \geq 0.$$

Note that $\Psi_{ii} > 1$, $1 \leq i \leq S$, so all terms are positive. Since $\frac{\Psi_{ki}}{\Psi_{ii}-1} \geq \frac{\Psi_{ki}}{\Psi_{ii}}$, the same argument as before is valid for the first term. For the second term, (38) implies that

$$\frac{\Psi_{i+1,i+1}}{\Psi_{i,i+1}} - \frac{\Psi_{i+1,i}}{\Psi_{i,i}-1} \geq 0, \qquad 1 \leq i \leq S-1,$$

and thereby that the expression inside the parenthesis is nonnegative for $k = i+1$. Further, since Ψ is TP2, $\frac{\Psi_{k,i+1}}{\Psi_{k,i}} \geq \frac{\Psi_{i+1,i+1}}{\Psi_{i+1,i}} \geq \frac{\Psi_{i,i+1}}{\Psi_{i,i}-1}$ for $k > i+1$, so $\frac{\Psi_{k,i+1}}{\Psi_{i,i+1}} \geq \frac{\Psi_{ki}}{\Psi_{ii}-1}$ for all $k \geq i+1$. Thus, the result follows. An identical argument shows that the result is also valid for $i = j+1$, so the result indeed holds for all i. We are done.

Proof of Corollary 1

When $S = 2$, only one condition for (38), $i = 1$, needs to be checked. The general form of Φ is

$$\Phi = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}, \qquad 0 \leq p, q \leq 1,$$

leading to

$$\Psi = \frac{1}{(1-\delta(1-p))(1-\delta(1-q)) - \delta^2 pq}\begin{bmatrix} 1-\delta(1-q) & \delta q \\ \delta p & 1-\delta(1-p) \end{bmatrix}.$$

The left hand side of (38) is now $\frac{1}{(1-\delta(1-p))(1-\delta(1-q))-\delta^2 pq}$, which is (weakly) greater than both Ψ_{11} and Ψ_{22}, so (38) is indeed satisfied. It is also greater than zero, which is the TP2 condition (37). We are done.

Proof of Proposition 8

The case $S = 2$ has already been covered. We therefore assume that $S \geq 3$. It follows that $R = \Psi^{-1} = (I - \delta\Phi)$ is a tridiagonal matrix with elements $R_{1,1} = R_{S,S} = z(1+q)$, $R_{i,i} = z(2+q)$, $i = 2, \ldots, S-1$, and $R_{i,i+1} = R_{i+1,i} = -z$, $i = 1, \ldots, S-1$, where $z = \delta p$, and $q = \frac{1-\delta}{\delta p}$. Now, following the argument in Pena (1995), since Ψ is a matrix with positive elements, it follows that R is a tridiagonal so-called

M-matrix, and that $\Psi = R^{-1}$ therefore is a totally positive matrix (of any order). Thus, (37) is satisfied. To show (38), we use the representation of Ψ in Meek (1980):

$$\begin{aligned}
\Psi_{i,j} &= e_{\min(i,j)} e_{S-\max(i,j)+1} + A_j e_{S-i+1} + B_j e_i, \\
e_i &= \frac{\sinh((i-1)\theta)}{\sinh((S-1)\theta)}, \\
\theta &= \operatorname{arcosh}\left(1 + \frac{q}{2}\right), \\
A_j &= \frac{1}{(1+q-e_{S-1})^2 - e_2^2}\left((1+q-e_{S-1})e_{S-j+1} + e_2 e_j\right), \\
B_j &= \frac{1}{(1+q-e_{S-1})^2 - e_2^2}\left((1+q-e_{S-1})e_j + e_2 e_{S-j+1}\right).
\end{aligned}$$

Some algebra leads to the following expression

$$\Psi_{i+1,i+1} - \frac{\Psi_{i,i+1}\Psi_{i+1,i}}{\Psi_{i,i}} = \frac{qe_2}{(1-\delta)\alpha}(e_{i+1} + e_{S-i}\beta)^2 \left(\frac{e_{S-i} + e_{i+1}\beta}{e_{i+1} + e_{S-i}\beta} - \frac{e_{S-i+1} + e_i\beta}{e_i + e_{S-i+1}\beta}\right), \text{ where}$$
$$\beta = \frac{1+q-e_{N-1}}{e_2}, \qquad \alpha = (1+q-e_{N-1})^2 - e_2^2.$$

If this expression is (weakly) greater than one, then $\Psi_{i+1,i+1}\Psi_{i,i} - \Psi_{i,i+1}\Psi_{i+1,i} \geq \Psi_{i,i}$. Now, $\delta = \frac{1}{1+pq} \geq \frac{1}{1+0.5q}$, so a sufficient condition is that

$$T \stackrel{\text{def}}{=} \frac{(2+q)e_2}{\alpha}(e_{i+1} + e_{S-i}\beta)^2 \left(\frac{e_{S-i} + e_{i+1}\beta}{e_{i+1} + e_{S-i}\beta} - \frac{e_{S-i+1} + e_i\beta}{e_i + e_{S-i+1}\beta}\right) \geq 1.$$

Plugging in $q = 2(1 + \cosh(\theta))$, and simplifying, we get

$$T = \frac{2e^{\theta}\left(e^{2i\theta} + e^{(1+2S)\theta}\right)\cosh(\theta)}{e^{2i\theta} + e^{(3+2S)\theta}}.$$

Taking the derivative with respect to θ yields

$$\frac{\partial T}{\partial \theta} = -\frac{2\left(e^{(4+4S)\theta} - e^{(2+4i)\theta}\right) + (2S+1-2i)\left(e^{(5+2i+2S)\theta} - e^{(1+2i+2S)\theta}\right)}{(e^{2i\theta} + e^{(3+2S)\theta})^2},$$

which by inspection is seen to be strictly negative. Thus, T is strictly decreasing in θ. Finally, an asymptotic expansion of T for large θ implies that

$$T = \frac{e^{4S\theta}(1 + o(\theta))}{e^{4S\theta}(1 + o(\theta))},$$

so $\lim_{\theta\to\infty} T(\theta) = 1$. Thus, $T(\theta) > 1$ for all θ, and therefore $\Psi_{i+1,i+1}\Psi_{i,i} - \Psi_{i,i+1}\Psi_{i+1,i} \geq \Psi_{i,i}$. An identical argument shows that $\Psi_{i+1,i+1}\Psi_{i,i} - \Psi_{i,i+1}\Psi_{i+1,i} \geq \Psi_{i+1,i+1}$, so (38) is satisfied. We are done.

Proof of Proposition 9

The proof is very similar to the proof of Proposition 7. From (40), it follows that

$$\mu_i = \frac{\sum_j \left[\Phi_{ij} + \sum_k \Phi_{ik}\frac{1}{m_k}\hat{\Psi}_{kj}m_j\right]\pi_j}{\sum_j \frac{1}{m_i}\hat{\Psi}_{ij}m_j\pi_j}, \tag{73}$$

where $\hat{\Psi} = \Psi - I$. Since $\hat{\Psi} = \Psi - I = \delta\Phi + \delta^2\Phi^2 + \cdots = \delta(\Phi + \Phi(\delta\Phi + \cdots)) = \delta(\Phi + \Phi\hat{\Psi})$ it follows that

$$\frac{\sum_j \left[\Phi_{ij} + \sum_k \Phi_{ik}\hat{\Psi}_{kj}\right]m_j\pi_j}{\sum_j \hat{\Psi}_{ij}m_j\pi_j} = \frac{1}{\delta},$$

which we use to rewrite

$$\mu_i = \frac{1}{\delta} + m_i \sum_k \left(r_i(k) \Phi_{ij} \frac{\sum_j \Psi_{kj} m_j \pi_j}{\sum_j \hat{\Psi}_{ij} m_j \pi_j} \right),$$

where, as before, $r_i(k) \stackrel{\text{def}}{=} \frac{1}{m_k} - \frac{1}{m_i}$, which defining $\xi_j \stackrel{\text{def}}{=} m_j \pi_j$, can be written

$$\mu_i = \frac{1}{\delta} + m_i \sum_k \left(r_i(k) \Phi_{ij} \frac{\sum_j \Psi_{kj} \xi_j}{\sum_j \hat{\Psi}_{ij} \xi_j} \right),$$

and since ξ_j is strictly increasing in α_j, we study

$$\frac{\partial \mu_i}{\partial \xi_n} = m_i \frac{\hat{\Psi}_{in}}{\sum_j \hat{\Psi}_{ij} \xi_j} \sum_k \left(r_i(k) \Phi_{ij} \left(\frac{\Psi_{kn}}{\hat{\Psi}_{in}} - \frac{\sum_j \Psi_{kj} \xi_j}{\sum_j \hat{\Psi}_{ij} \xi_j} \right) \right),$$

We define

$$F_n = \sum_{k<i} \left(r_i(k) \Phi_{ij} \left(\frac{\Psi_{kn}}{\hat{\Psi}_{in}} - \frac{\sum_j \Psi_{kj} \xi_j}{\sum_j \hat{\Psi}_{ij} \xi_j} \right) \right),$$
$$G_n = \sum_{k>i} \left(r_i(k) \Phi_{ij} \left(\frac{\Psi_{kn}}{\hat{\Psi}_{in}} - \frac{\sum_j \Psi_{kj} \xi_j}{\sum_j \hat{\Psi}_{ij} \xi_j} \right) \right),$$

to get

$$\frac{\partial \mu_i}{\partial \xi_n} = m_i \frac{\hat{\Psi}_{in}}{\sum_j \hat{\Psi}_{ij} \xi_j} (F_n + G_n).$$

A similar argument as in the proof of Proposition 7, shows that $F_{n+1} \le F_n$, and $G_{n+1} \ge G_n$ $(n < S)$ so it is indeed that case that if $\frac{\partial \mu_i}{\partial \xi_n} \ge 0$, then $\frac{\partial \mu_i}{\partial \xi_{n+1}} \ge 0$. Similarly, if $\frac{\partial \mu_i}{\partial \xi_n} \le 0$ $(n > 1)$, then $\frac{\partial \mu_i}{\partial \xi_{n-1}} \le 0$.

It remains to show that both the case $\frac{\partial \mu_i}{\partial \xi_n} \ge 0$, and $\frac{\partial \mu_i}{\partial \xi_n} \le 0$ occur, but this follows easily from the fact that

$$\min_j \frac{a_j}{b_j} \le \frac{\sum_j a_j \xi_j}{\sum_j b_j \xi_j} \le \max_j \frac{a_j}{b_j},$$

for arbitrary positive a_j ξ_j and strictly positive b_j. We are done.

References

Abreu, D., 1988, On the theory of infinitely repeated games with discounting, *Econometrica* 56, 383–96.

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi, 2011, The network origins of aggregate fluctuations, *Econometrica* forthcoming.

Bagwell, K., and R. Staiger, 1997, Collusion over the business cycle, *RAND Journal of Economics* 28, 82–106.

Dixit, A. K., and J. E. Stiglitz, 1977, Monopolistic competition and optimum product diversity, *American Economic Review* 67, 297–308.

Gabaix, X., 2011, The granular origins of aggregate fluctuations, *Econometrica* 79, 733–772.

Hoberg, G., and G. Phillips, 2010, Real and financial industry booms and busts, *Journal of Finance* 65, 45–86.

Hou, K., and D. Robinson, 2006, Industry concentration and average stock returns, *Journal of Finance* 61, 1927–1956.

Jovanovic, B., 1987, Micro shocks and aggregate risk, *Quarterly Journal of Economics* 102, 395–409.

Karlin, S., 1968, *Total Positivity* (Stanford University Press, Stanford, CA).

Kydland, F. E., and E. C. Prescott, 1982, Time to build and aggregate fluctuations, *Econometrica* 50, 1345–1370.

Long, B., and C. I. Possner, 1983, Real business cycles, *Journal of Political Economy* 91, 39–69.

Lyandres, E., and M. Watanabe, 2010, Product market competition and equity returns, Technical report, Boston University, Unpublished manuscript.

Meek, D., 1980, The inverses of some matrices deviating slightly from a symmetric, tridiagonal, toeplitz form, *SIAM Journal of Numerical Analysis* 17, 39–43.

Novy-Marx, R., 2010, The other side of value: Good growthand the gross profitability premium, Technical report, University of Chicago, Unpublished manuscript.

Opp, M. M., 2010, Tariff wars in the ricardian model with a continuum of goods, *Journal of International Economics* 80, 212–225.

Parlour, C., R. Stanton, and J. Walden, 2011, Revisiting asset pricing puzzles in an exchange economy, *Review of Financial Studies* 24, 629–674.

Pena, J. M., 1995, m-matrices whose inverses are totally positive, *Linear Algebra and its Applications* 221, 189–193.

Ravn, M., S. Schmitt-Grohe, and M. Uribe, 2006, Deep habits, *Review of Economic Studies* 73, 195–218.

Rotemberg, J. J., and G. Saloner, 1986, A supergame-theoretic model of price wars during booms, *American Economic Review* 76, 390–407.

Rotemberg, J. J., and M. Woodford, 1992, Oligopolistic pricing and the effects of aggregate demand on economic activity, *Journal of Political Economy* 100, 1153–1207.

van Binsbergen, J. H., 2007, Deep Habits and the Cross Section of Expected Returns, *SSRN eLibrary* .