Capital Investment, Equity Returns and Aggregate Dynamics in Oligopolistic Production Economies*

Hitesh Doshi
University of Houston

Praveen Kumar
University of Houston

This Version: January 2024

Abstract

We analyze the effects of tacit collusion in a dynamic general equilibrium model of oligopolistic sectors with capital investment, real frictions, and entry. Through their effects on equilibrium- and off-equilibrium stock prices, fundamental shocks impact incentives for defection from tacit collusion, amplifying the interaction between the real economy and financial markets as well as firms' risk exposure. Quantitively, the model's endogenous firm- and aggregate-level dynamics help explain real and financial data. We also find quantitative and empirical support for novel theoretical predictions regarding the effects of firm, industry and macrofinance characteristics on cross-sectional and time-varying relation of concentration and returns.

*We thank four anonymous referees and Itay Goldstein (editor) for very helpful comments and guidance. For helpful comments and discussions we also thank Franklin Allen, Stephen Arbogast, Cristina Arellano, Tim Bresnahan, Laurent Calvert, Adlai Fisher, Lorenzo Garlappi, Tom George, Nils Gottfries, Kris Jacobs, Ravi Jagannathan, Mamoud Medhat, Elisa Pazaj, Valerie Ramey, Enrique Schroth, Sang Seo, Ken Singleton, Roberto Steri, Jincheng Tong, Vijay Yerramilli, participants in seminars at various universities, the Annual Meeting of the American Economic Association, the Corporate Policies and Asset Pricing (COAP) Annual Conference, the North American Summer Meeting of the Econometric Society, and the Northern Finance Association Conference. Send correspondence to: Praveen Kumar, C.T. Bauer College of Business, 4750 Martin Luther King Blvd, Houston, TX 77204; +1(832)-452-5625; e-mail: pkumar@uh.edu.

1 Introduction

Oligopolistic industries are ubiquitous and strategic behavior of firms in oligopolies receives long-standing interest.[1] In particular, the literature highlights the role of production capacity in protecting oligopoly profits through tacit collusion (Brock and Scheinkman 1985; Benoit and Krishna 1987) and entry deterrence (Spence 1977; Dixit 1980). A related literature examines the effects of oligopolistic collusion on aggregate real outcomes (Rotemberg and Woodford 1992; Opp et al. 2014) and risk factors (Dou et al. 2021). Meanwhile, another burgeoning literature considers the effects of industry concentration and entry on equity returns. However, the effects of capital investment and strategic capacity utilization on firm-level equity returns and aggregate dynamics through the channels of tacit collusion, entry deterrence and endogenous capacity depreciation are still largely unexplored. In this paper, we attempt to fill this gap.

The quantitative analysis of returns and aggregate dynamics with tacit collusion in production economies with investment remains open because the existing literature mainly examines oligopolistic outcomes in endowment economies (Dou et al. 2021) or in production economies without investment (Opp et al. 2014; Loualiche 2021). Meanwhile, there are conflicting empirical results on the relation of industry concentration and returns: One strand of the literature documents a negative relationship (Hou and Robinson 2006; Gu 2016); while another strand presents a positive relationship (Bustamante and Donangelo 2017; Corhay et al. 2020) or finds no significant relationship (Ali et al. 2013). Our analysis and results help address both these quantitative and empirical issues.

We develop a dynamic general equilibrium model of a production economy with capital investment, composed of multiple oligopolistic consumer goods industries (or sectors) with unlevered firms. In concentrated sectors, there is tacit collusion on product prices through strategic capacity utilization (production) and investment along subgame perfect equilibrium (SPE) paths, enforced by credible threats of "punishment" through Bertrand-Nash price competition following defection. Analytically, the novel aspect of the model is equilibrium characterization of firms' real decisions and equity returns, along with the resultant endogenous aggregate dynamics, through a dynamic optimization problem of maximizing equity values subject to a tacit collusion incentive compatibility constraint, namely, that current equity gains from defection not exceed losses from defection.

[1] Observations on strategic interaction among oligopolistic firms occur at least as far back as Smith (1776). While the early literature (Cournot 1838) analyzed static interactions, the literature in the past few decades focuses on tacit collusion in dynamic oligopolies. Friedman (1983), Bresnahan (1989), Feuerstein (2005), and Green et al. (2013) provide useful surveys of these literatures.

Focusing on oligopolistic strategic interaction with capital investment provides novel insights on the interaction between the real economy and financial markets at both firm and aggregate levels. Fundamental (productivity) shocks impact incentives for defection from tacit collusion through their effects on equilibrium- and off-equilibrium stock prices. In response, firms adjust their real policies to maintain tacit collusion, thereby linking financial markets and real decisions. In general equilibrium, tacit collusion mediates the effects of sectoral shocks on endogenous aggregate dynamics in income, price index, consumption and the risk-free rate. Indeed, productivity shocks in low-concentration sectors have significant effects on tacit-collusive real and financial outcomes in concentrated sectors because of *intratemporal* elasticity of substitution between industry products.

We find that these real and financial interactions in concentrated oligopolistic sectors are quantitatively significant and better explain firm and aggregate real and financial data compared to sectors with Bertrand-Nash price competition. Thus, while productivity shocks play a central role in the real business cycle and macrofinance models in competitive settings, these shocks drive micro and aggregate outcomes in oligopolistic production economies by affecting firm-level defection incentives channeled through equilibrium- and off-equilibrium equity prices.[2]

It also follows that firms' risk exposure to fundamental shocks is amplified by the possibility of defection from tacit collusion and off-equilibrium behavior plays a significant role in this exposure. Indeed, our analysis provides a novel perspective on the role of capital and real frictions— that is, fixed and capital adjustment costs, endogenous depreciation, and entry costs—on priced risks. In the existing literature, real frictions amplify firms' risk exposure by restricting responses to fundamental shocks (Jermann 1998; Carlson et al. 2004). Our framework highlights an additional channel: On- and off-equilibrium industry production capacities are central to tacit collusion because they determine short-term benefits and long-term expected costs of defection. Through their influence on capital investment and entry, real frictions affect defection incentives and hence priced risks .

Similar to the literature (Baxter and Curcini 1993; Boldrin et al. 2001), we quantitatively analyze a two-sector economy. In our case, these sectors comprise of a tacit-collusive concentrated sector and a benchmark low-concentration sector with Bertrand-Nash price competition. The representative consumer has Epstein and Zin (1989) preferences. We assume a moderate risk aversion of 5 and calibrate the intertemporal elasticity of substitution consistent with recent production-based asset pricing literature (e.g., Croce 2014). Our modeling of sectoral productivity shocks—the exogenous

[2]The significant effect of off-equilibrium equity values on equilibrium outcomes is consistent with the well known importance of off-equilibrium behavior on subgame perfect equilibrium paths (Selten 1975; Kreps and Wilson 1982).

sources of risk in our model—is parsimonious: They follow an AR(1) process in logs and we adopt a conservative calibration of the second moments of these shocks.

To isolate the effects of entry and exit, we first analyze a baseline model with fixed concentration. The baseline model generates volatility of aggregate real consumption growth and real risk-free rate that match the data. But the volatilities of log changes in aggregate (financial) income, price index and the risk-free rate understate the data. At the firm level, for the concentrated sector, the baseline model generates an annual equity premium and its volatility that are higher than benchmark production-based asset pricing models but lower than in the data. On the real side, the baseline model understates the volatilities of capacity utilization and log changes in capital investment (for the concentrated sector) relative to the data. However, the empirical performance of the tacit collusion equilibrium path is substantially better relative to the low-concentration sector. The baseline model does not match well the autocorrelation of aggregate outcomes, but is reasonably close to the data in terms of autocorrelation of firm-level investment. The model is also largely consistent with simultaneous correlations of real variables in the data.

We find that the interaction of endogenous entry in the tacit-collusive sector and capital adds significant volatility to equilibrium real and financial outcomes. At the firm level, the volatilities of capacity utilization and investment increase significantly with entry and no longer understate the data. Similarly, at the aggregate level, the volatilities of income, price index, consumption and risk-free rate also rise significantly. Notably, entry significantly improves the model performance with respect to the autocorrelations of aggregate consumption and income. Endogenous changes in industry structure through entry in concentrated oligopolies, along with endogenous capital and real frictions, therefore appear important in explaining aggregate dynamics.

There is an intuition that with higher sectoral and aggregate volatility, endogenous entry should further amplify firms' risk exposure to fundamental shocks in the presence of tacit collusion and capital. Consistent with this, relative to the baseline, the model with entry generates a significantly greater equity premium with high volatility. Moreover, higher capital adjustment costs and volatility of sectoral productivity shocks in the competitive sector significantly amplify the effects of entry on firms' risk exposure in industries with tacit collusion.

The presence of tacit collusion also affects the link between micro characteristics and aggregate dynamics. The existing literature generally finds that variations in micro-level adjustment costs have negligible effects on aggregate dynamics because of compensating changes in prices (Veracierto 2002;

Khan and Thomas 2008). In contrast, we find significant aggregate effects of variations in firm-level adjustment costs and productivity processes in the presence of tacit-collusive prices and real frictions.

Our analysis also shows that capital and real frictions in tacit-collusive oligopolies help explain the conflicting results in the literature on the relation of industry concentration and equity returns. Specifically, higher concentration has two opposing effects on markups and, hence, equity payoffs. There is an unambiguous negative "market dilution" effect of lower concentration because, for a given industry capacity, firms' profits are positively related to concentration. This dilution effect ceteris paribus leads to a positive relation of concentration and payoffs. On the other hand, lower industry concentration has a positive "punishment" effect on equilibrium payoffs by facilitating a credible threat of substantially lower equity values following defections. When the dilution effect dominates, lower concentration is ceteris paribus negatively related to returns, and conversely when the punishment effect is relatively strong.

We build on this analysis to generate unambiguous predictions regarding the effects of operating leverage, operating profitability, and investment on the relation of concentration and returns. To illustrate, equity values of firms with high fixed or adjustment costs are relatively low during recessions, so that the punishment effect becomes weak for such firms in high SDF states. Hence, higher fixed or adjustment costs should negatively affect the relation of concentration and returns. An opposing argument applies for high operating profitability firms and the concentration-returns relation should switch in sign for such firms. We quantitatively confirm these predictions in an industry equilibrium setting. Meanwhile, higher depreciation costs of capacity utilization (Greenwood et al. 1988; Jaimovich and Rebelo 2009) ceteris paribus lower optimal production and raise industry prices, weakening the punishment effect. Therefore, the concentration-returns relation is predicted to be more negative in industries with higher marginal depreciation costs of capacity utilization. Moreover, incumbents in high entry-threat industries with scale economies can deter entry through large capacities (Spence 1977; Dixit 1981), strengthening the link between concentration and capital investment, as well as the effect of investment on the concentration-returns relation.

We also examine time-variation in the effects of concentration on returns. Duffee (2005) finds that the conditional covariance between the SDF and (aggregate) returns rises in high aggregate return periods, suggesting that the concentration-returns relation will ceteris paribus be greater in high aggregate return states. Moreover, a growing literature shows that sales concentration and markups have increased in recent decades for a few "super efficient" firms, accompanied by technological change that raises the ratio of fixed-to-marginal costs (Berry et al. 2019; Autor et al. 2020). A secular trend of increasing fixed-to-marginal cost ratios should further weaken the punishment effect for high fixed cost firms and strengthen the negative effect of operating leverage on the concentration-return relation. In contrast, superior efficiency of larger incumbents reduces reliance on capital investment for entry deterrence in high entry-threat industries, weakening the link between concentration and investment, as well as the effect of higher investment on the concentration-returns relation.

Our empirical tests utilize the *Text-based Network Industry Classification* (TNIC) Herfindahl-Hirschman Index of Hoberg and Phillips (2016) as the concentration measure because the TNIC better matches our model. Our base sample includes all non-financial Compustat firms during 1989-2019 with available concentration measure. Our cross-sectional tests utilize Fama and MacBeth (1973) regressions on firms' stock returns, controlling for Fama and French (1992, 1993) factors and leverage. We find empirical support for the above-mentioned predictions from the model.

Related literature The interaction of real economy and financial markets attracts much attention (Morck et al. 1990).[3] The existing literature highlights the real effects of financial markets based on the information role of security prices (Chen et al. 2007; Bond et al. 2012). We present a novel channel for the real effects of stock prices in dynamic oligopolies that are widely prevalent. Indeed, we find that both equilibrium and off-equilibrium equity prices have an empirically significant effect on firm-level real decisions as well as on aggregate dynamics.

A growing literature examines real and financial outcomes in industries with imperfect competition and entry. Some papers in this literature—such as, Aguerrevere (2009), Morrellec and Zhdanov (2019), Bustamente (2015), Farhi and Guorio (2018) and Corhay et al. (2020)—consider investment, but not the interaction of capital and real frictions with tacit collusion in a dynamic general equilibrium setting, as is the case with our study. Opp et al. (2014) and Dou et al. (2021) analyze the effects of tacit collusion in general equilibrium models with oligopolistic sectors. However, Opp et al. (2014) consider a production economy with only labor input and hence do not consider the role of capital, while Dou et al. (2021) analyze an endowment economy model with exogenous aggregate output. Similarly, other papers in the literature, such as Bustamente and Donangelo (2017) do not consider endogenous investment. Corhay et al. (2020) and Loualiche (2021) positively link entry threat and risk premium in imperfectly competitive industries but do not consider the interaction of tacit collusion with capital and real frictions.

[3] Morck et al. (1990) frame this issue by asking whether the stock market is a "sideshow."

In contrast to this literature, we focus on the effects of tacit collusion on firm- and aggregate-level real and financial variables in dynamic production economies with capital investment and entry. Our study contributes by theoretically, quantitatively, and empirically showing the importance of tacit collusion with endogenous capital and entry in helping explain firm-level and aggregate data on salient real and financial variables. In particular, we highlight the importance of off-equilibrium capital and real frictions on firms' priced risks and real behavior, and the significant economy-wide effects of fundamental shocks through the tacit collusion channel in oligopolistic sectors. Our study is also novel in highlighting the role of real frictions, strategic entry deterrence, and time-varying risks in helping resolve the ambiguous relation of concentration and returns in the literature.

The literature also considers effects of operating leverage (Lev 1974; Novy-Marx 2011), operating profitability and investment (Aharoni et al. 2013; Fama and French 2015), and entry threat (Corhay et al 2020; Loualiche 2021) on the cross-section of returns, as well as the time-varying effects of market returns (Duffee 2005). In contrast, these firm, industry and macrofinance characteristics play a prominent role in our analysis because they help resolve the theoretically ambiguous effect of industry concentration on equity returns. More generally, these conflicting effects point to the importance of capital and real frictions in understanding the cross-sectional relation of concentration and returns.

The received dynamic oligopoly literature also focuses on symmetric SPE that maximize profits through worst-possible credible punishment following defection (Abreu 1986) and notes the theoretically ambiguous effect of concentration on markups (Brock and Scheinkman 1985). In particular, Rotemberg and Woodford (1992) use this framework to analyze the effects of aggregate demand shocks on economic activity in an oligopolistic production economy with tacit collusion. However, our analysis is distinct from this literature because of our emphasis on the interaction of tacit collusion, capital investment, entry and entry deterrence, and real frictions with equity markets; our focus yields novel results on the interaction of financial markets with the real economy, as well as on the relation of industry concentration and equity returns.

2 The basic model

We consider an infinite horizon, discrete time model (indexed by $t = 1, ...$). For expositional ease, we initially describe and analyze the basic model, which we subsequently extend to allow entry and exit, as well as endogenous depreciation and capacity-based entry deterrence.

The economy consists of J consumer goods sectors (or industries) that each produce a homogenous good, utilizing capital input from a (numeraire) investment goods sector. Firms in each industry are unlevered and their equity shares are traded in frictionless security markets. Cash flows, determined by operating profits net of investment and fixed capacity costs, are distributed to shareholders as dividends. There is a continuum of identical consumer-investors (CIs) in the economy; the number of CIs is normalized to unity, without loss of generality. CIs purchase consumption goods and channel their savings back to firms through equity investment or investment in a riskless asset. Firms maximize discounted expected shareholder value of real dividends.

2.1 Consumer preferences and consumption-investment choice

At each t, the representative CI chooses the consumption profile $\mathbf{c}_t = (c_{1t}, ...c_{Jt})$ with given prices $\mathbf{p}_t = (p_{1t}, ..., p_{Jt})$, subject to an income constraint. The CI can invest in a riskless security (f) with a unit mass that makes a unit payment next period. Asset holdings at the beginning of t are denoted by the vector $\mathbf{q}_t$ with the associated dividend vector $\mathbf{d}_t$ (and $d_{ft} = 1$). Along with consumption, the CI simultaneously chooses new asset holdings $\mathbf{q}_{t+1}$ taking as given ex-dividend equity prices $\mathbf{s}_t$.

The CI has the constant elasticity of substitution (CES) form of Kreps and Porteus (1978) preferences (Epstein and Zin 1989). At each t, the lifetime utility of the CI is recursively given by

$$\mathcal{U}_t = \left[(1-\alpha)C_t^{1-\eta} + \alpha\mathbb{E}_t\left[\mathcal{U}_{t+1}^{1-\gamma}\right]^{\frac{1-\eta}{1-\gamma}}\right]^{\frac{1}{1-\eta}}, \tag{1}$$

where $C_t \equiv \left[\sum_{j=1}^{J}\phi_j(c_{jt})^{(\sigma-1)/\sigma}\right]^{\sigma/(\sigma-1)}$ is the aggregated CES consumption basket; $\sigma > 1$ is the common elasticity of substitution (ES); and $0 < \phi_j < 1$ are aggregation weights. Furthermore, α is the subjective rate of impatience; γ determines the degree of risk aversion; and η^{-1} measures the IES. For tractability, we do not consider non-financial income. The CI's budget constraint is thus

$$\mathbf{p}\cdot\mathbf{c}_t \le \mathbf{q}_t\cdot(\mathbf{d}_t + \mathbf{s}_t) - \mathbf{q}_{t+1}\cdot\mathbf{s}_t \equiv W_t, \tag{2}$$

where W_t is the net financial income. Because preferences are strictly increasing, the budget constraint (2) will be binding in optimum. Intraperiod optimization yields the consumption demand functions (Online appendix A.1):

$$c_{jt}(\mathbf{p}_t, W_t) = \frac{W_t}{P_t}\left[\frac{P_t\phi_j}{p_{jt}}\right]^{\sigma},\ j = 1, ..J, \tag{3}$$

where $P_t \equiv \left[\sum_{j=1}^{J}(\phi_j)^{\sigma}(p_{jt})^{1-\sigma}\right]^{1/(1-\sigma)}$ is the aggregate (consumer) price index. Consumer optimum implies that aggregate real consumption $C_t = \frac{W_t}{P_t}$.

The CI's portfolio optimization equates real current security prices $\frac{\mathbf{s}_t}{P_t}$ to expected present value of real equity payoffs next period. Letting $\Lambda_t \equiv \frac{\partial \mathcal{U}_t}{\partial \mathcal{C}_t} = (1-\alpha)C_t^{-\eta}\mathcal{U}_t^{\eta}$ denote the marginal valuation at t, the SDF for the one-period investment horizon is $\Lambda_{t,t+1} \equiv \frac{\Lambda_{t+1}}{\Lambda_t}$, where

$$\Lambda_{t,t+1} = \alpha\left(\frac{C_{t+1}}{C_t}\right)^{-\eta}\left[\frac{\mathcal{U}_{t+1}}{\mathbb{E}_t\left[\mathcal{U}_{t+1}^{1-\gamma}\right]^{1/(1-\gamma)}}\right]^{\eta-\gamma}. \tag{4}$$

Hence, asset prices satisfy

$$\frac{\mathbf{s}_t}{P_t} = \mathbb{E}_t\left[\Lambda_{t,t+1}\left(\frac{\mathbf{d}_{t+1}+\mathbf{s}_{t+1}}{P_{t+1}}\right)\right]. \tag{5}$$

It will be useful to express (5) in terms of nominal returns,[4] namely, the gross return $R_{t+1} = \left(\frac{d_{t+1}+s_{t+1}}{s_t}\right)$ (with $R_{f,t+1} = (1/s_{ft})$) by defining the *nominal* SDF $\Omega_{t,t+\tau} \equiv P_t\left(\frac{\Lambda_{t,t+\tau}}{P_{t+\tau}}\right), \tau = 1, 2, ...$ We can thus rewrite (5) as the equilibrium asset market condition $\mathbf{1} = \mathbb{E}_t\left[\Omega_{t,t+1}\mathbf{R}_{t+1}\right]$, where $\mathbf{1}$ and $\mathbf{R}_{t+1}$ are the unit and gross nominal returns vectors, respectively.

2.2 Production and capital investment

The representative firm in the typical sector produces output Y_t through the production function[5]

$$Y_t = A_t(u_t K_t)^{\psi}, \tag{6}$$

where K_t is the firm's capital stock at the beginning of t; A_t represents the stochastically evolving industry-wide productivity level; $0 < \psi < 1$ is the output elasticity of capital; and $u_t \in [0,1]$ is the firm's capacity utilization at t. The sectoral productivity shocks follow an AR(1) process in logs,

$$a_t = \rho a_{t-1} + \varepsilon_t, \tag{7}$$

where $a_t = \ln(A_t)$ and ε_t are mean zero variables with a stationary variance-covariance matrix $[\lambda_{ij}]$ across sectors. The sectoral productivity shocks are the only exogenous sources of risk in the economy.

[4] The literature on cross-sectional and time-series analysis of expected returns usually utilizes nominal returns (Fama and French 1992, 1993; Lewellen 1999). Our cross-sectional and time-series tests also use nominal stock returns.

[5] For notational ease, we suppress subscripts for sectors and firms unless necessary for exposition.

Firms can undertake capital investment, I_t, by deploying a capital good and production capacity evolves as

$$K_{t+1} = (1-\delta)K_t + I_t, \tag{8}$$

where $\delta \in (0,1)$ is the depreciation rate. For tractability, it is assumed that there is an infinitely elastic supply of the capital good with unit price, so that relative prices for consumption goods are defined in terms of the capital good.[6] Firms are subject to strictly convex capacity adjustment costs (Lucas 1967; Liu et al. 2009) and the investment cost function is

$$\Psi(I_t, K_t) = I_t + 0.5\varphi\left(\frac{I_t}{K_t} - \delta\right)^2 K_t, \tag{9}$$

where φ is the capacity adjustment cost. The quadratic investment cost specification is common in the literature (Summers 1981; Liu et al. 2009). Consistent with the real business cycle (RBC) literature (Chari et al. 2000), we assume there are no adjustment costs in the steady state.

2.3 Equilibrium paths

All functional relationships in the model are common knowledge, as is the state of the economy, denoted by the sequential filtration Γ_t, which includes all economically relevant variables at the beginning of t. Because industry productivity shocks are the only exogenous shocks in the model, uncertainty regarding exogenous events in period t is resolved conditional on Γ_t.

Given Γ_t, firms (denoted by n) in each industry independently choose their prices $\mathbf{p}_t = (p_{nt})_{n=1}^N$, where N is the number of firms in the industry. Firm-level residual demand functions $F_{nt}(p_{nt}; \Gamma_t, \mathbf{p}_t)$ are determined similar to the literature (Kreps and Scheinkman 1983; Brock and Scheinkman 1985): The lowest-price firms sell as much of the quantity demanded subject to their production capacity and all firms divide the demand equally. There are sector-specific constant variable costs h so that operating profits are $\Pi_{nt}(p_{nt}; p_t) = (p_{nt} - h)F_{nt}(p_{nt})$, while capacity utilization is $u_{nt} = \min\left(1, \frac{F_{nt}}{A_t(K_{nt})^\psi}\right)$. Firms then independently decide their investments with the resultant dividends

$$D_t = \Pi_t - \Psi(I_t, K_t) - mK_t, \tag{10}$$

[6] In the literature, multisector competitive production economy models with investment consider heterogeneous capital input prices (Horvath 2000). Our modeling approach facilitates focus on tacit collusion in consumer goods prices, as in the received oligopoly literature. It is also useful in highlighting the effects of tacit collusion because Euler conditions in our model can be readily compared with the RBC or macrofinance literature. In addition, this approach allows a parsimonious representation of fundamental shocks, which is helpful in our quantitative analysis.

where m is the (sector-specific) fixed cost for operating production capacity. Hence, $d_t = \frac{D_t}{Q}$. Negative dividends are financed through equity issuance.

Firms are instructed by shareholders to maximize the conditional present value of real dividends, $\left[\sum_{\tau=0}^{\infty} \Lambda_{t,t+\tau}\left(\frac{D_{t+\tau}}{P_{t+\tau}}\right)\right]$; or, in terms of the nominal SDF $\Omega_{t,t+\tau} \equiv P_t\left(\frac{\Lambda_{t,t+\tau}}{P_{t+\tau}}\right)$, to maximize $\mathbb{E}_t\left[\sum_{\tau=0}^{\infty} \Omega_{t,t+\tau} D_{t+\tau}\right]$. Using the Bellman representation, we can define the nominal *cum-dividend* value function recursively by $V_t(\Gamma_t) = \max_{(p_t,I_t)} D_t + \mathbb{E}_t\left[V_{t+1}(\Gamma_{t+1})\right]$. The ex-dividend value of the firm is denoted by S_t. Finally, similar to the literature, we adopt the convention that while there is imperfect competition within sectors, firms take aggregate quantities and the SDF as given.

Our model specifies a multi-stage game with perfectly observed actions (Fudenberg and Tirole 1991), where Γ_t defines a subgame. Along a SPE, at every Γ_t : (i) the CI chooses optimal consumption and portfolio policies (see (3) and (5)); (ii) in each sector, the profile of firms' price and investment strategies (p_t, I_t) comprise a Nash equilibrium; (iii) no firm strictly gains from deviating from the prescribed action at any stage and then returning to the equilibrium path thereafter; (iv) prices p_t clear the product market in each sector; and (v) stock prices $\mathbf{s}_t$ clear equity markets, that is, $\mathbf{q}_{t+1} = \mathbf{Q}_{t+1}$. Without loss of generality, firms' shares are normalized to unity at every t. Note that any SPE, $\langle\{p_{jt}, I_{jt}\}_{t=0}^{\infty}\rangle_{j=1}^{J}$, determines at each t firms' dividends across industries (see (10)) and hence aggregate income W_t (see (2)), price index P_t and real consumption $C_t = \frac{W_t}{P_t}$.

3 Tacit collusion and equilibrium characterization

We now characterize the equilibrium, focusing on the role of tacit collusion and capital investment. The consumption function (3) yields the pricing or inverse demand function

$$p_t(Y_t^{ind}) = \phi\left(Y_t^{ind}\right)^{-1/\sigma}\left(W_t P_t^{\sigma-1}\right)^{1/\sigma}. \tag{11}$$

Note that p_t is strictly decreasing in industry output and is well defined for all positive output levels. Hence, price is ceteris paribus negatively related to industry capacity utilization (see (6)).

It facilitates intuition to consider tacit collusion and defection incentives in a one-shot setting, with industry capacity K^{ind}. As usual, we can characterize the maximal industry (monopoly) profits in terms of prices or quantities. Using (3) and solving $\max_p\ c(\mathbf{p}^c, W)\,(p-h)$ yields the monopoly price $\hat{p} = \left(\frac{\sigma}{\sigma-1}\right)h$. Utilizing (11) and (6) then gives the *industry* profit-maximizing capacity utilization

$\hat{u} = \min\left(1, \left[\frac{(\sigma-1)W^{\frac{1}{\sigma}}P^{\frac{\sigma-1}{\sigma}}\phi}{\sigma h(A)^{\frac{1}{\sigma}}(K^{ind})^{\frac{\psi}{\sigma}}}\right]^{\frac{\sigma}{\psi}}\right)$. But one can reframe the choice problem in terms of choosing the profit maximizing industry capacity utilization, that is, $\hat{u} \in \arg\max_u (p(Y^{ind}) - h)A(uK^{ind})^{\psi}$, which yields the same solution. Now, $\hat{u}$ can be implemented by setting *firm-level* capacity utilization $\hat{u}_n = \hat{u}(N)^{\frac{\psi-1}{\psi}}$, which is decreasing in N (since $\psi < 1$). Hence, $\hat{u} > \hat{u}_n$ whenever $N \geq 2$.[7] But $\hat{u}_n < 1$ is not incentive compatible in the one-shot game: Any firm n can reduce its price slightly from $\hat{p}$, raise capacity utilization to 1 and increase profits by $(\hat{p} - h)A(K)^{\psi}(1 - (\hat{u})^{\psi})$.

With an infinite horizon, firms can attempt to enforce tacit collusion on prices—through strategic capacity utilization—based on credible retaliation threats. In particular, in the stochastic production and financial market setting of our model, defection incentives will evolve stochastically as firms evaluate their current profit gains from defection against loss of future equity value from retaliation. We turn now to characterize the equilibrium path.

3.1 Equilibrium characterization

The worst credible punishment path following a defection is the Bertrand-Nash (BN) SPE, where firms compete on prices subject to their capacity constraints. Consistent with the literature (Abreu 1986; Rotemberg and Saloner 1986), and to generate empirical predictions, we focus on the (symmetric) optimal SPE that maximizes equity values of industry firms through tacit collusion enforced by threats of switching to the BN SPE following defection. Along the BN SPE, firms choose capacity utilization and investment policies $(\tilde{u}_t, \tilde{I}_t)$ subject to the constraint of no gains from defection to a lower price. Because of product homogeneity and symmetric cost structure, the BN SPE involves full capacity utilization (Kreps and Scheinkman 1983). We indeed verify that $\tilde{u}_t = 1$ in our setting (Online appendix B.1). $\tilde{I}_t$ is determined by the Euler condition, $1 + \varphi\left(\frac{\tilde{I}_t}{K_t} - \delta\right) = \mathbb{E}_t\left[\frac{\partial \tilde{V}_{t+1}}{\partial I_t}\right]$ that equates marginal investment costs with marginal expected present value of gains

$$\mathbb{E}_t\left[\frac{\partial \tilde{V}_{t+1}}{\partial I_t}\right] = \mathbb{E}_t\left[\Omega_{t,t+1}\left\{\frac{\partial \tilde{\Pi}_{t+1}}{\partial K_{t+1}} - m + 0.5\varphi\left[\left(\frac{\tilde{I}_{t+1}}{K_{t+1}}\right)^2 - \delta^2\right] + (1-\delta)\left(1 + \varphi\left(\frac{\tilde{I}_{t+1}}{K_{t+1}} - \delta\right)\right)\right\}\right], \quad (12)$$

[7] Note that the industry output with $\hat{u}$ is $\hat{Y} = A(\hat{u}(NK))^{\psi}$, which is implemented by $\hat{u}_n$ such that $NA(\hat{u}_n K)^{\psi} = \hat{Y}$. Thus, $\hat{u}_n = \hat{u}(N)^{\frac{\psi-1}{\psi}}$. Furthermore, $\frac{1-\psi}{\psi}$ is a decreasing and convex function with the range $(0, \infty)$ on the domain $\psi \in (0, 1)$. Hence, $N^{-\frac{1-\psi}{\psi}} < 1$ and is decreasing in N.

where $\tilde{\Pi}_{t+1}$ are the operating profits with full capacity production by all firms (Online appendix B.2). The resultant dividends are $\tilde{D}_t = \tilde{\Pi}_{t+1} - \Psi(\tilde{I}_t, K_t) - mK_t$, which determine cum- and ex-dividend equity values $\tilde{V}_t$, $\tilde{S}_t$, respectively, through the asset market condition $\mathbb{E}_t\left[\Omega_{t,t+1}\left(\frac{\tilde{D}_{t+1}+\tilde{S}_{t+1}}{\tilde{S}_t}\right)\right] = 1$. In particular, the OEP BN equity values will significantly affect equilibrium real and financial variables.

We now characterize the equilibrium path for the representative firm in the typical sector, $\{u_t^*, I_t^*, S_t^*\}_{t=0}^{\infty}$. For any Γ_t, the product price is $p_t^* = p_t(Y_t^{ind*})$, with the industry output $Y_t^{ind*} = NA_t(u_t^* K_t^*)^{\psi}$, and resultant dividends D_t^* (see (10)). Defection potentially occurs by some firm choosing the price and capacity utilization $(p_t^* - \varepsilon, \bar{u}_t^*)$, where ε is arbitrarily small and $\bar{u}_t^*$ is the optimal defecting capacity utilization. But a firm willing to defect and suffer future loss of value from retaliation will seek to maximize current profits and set $\bar{u}_t^* = 1$ (see Online appendix C.1),[8] thus yielding the defection profits, $\bar{\Pi}_t^* = (p_t^* - h)A_t(K_t^*)^{\psi}$. Because a defecting firm reverts to the equilibrium path at the next (investment) stage, it invests I_t^* and the defecting dividends are $\bar{D}_t^* = \bar{\Pi}_t^* - \Psi(I_t^*, K_t^*) - mK_t^*$.

Following defection at any t, all firms switch to the punishment (BN) equilibrium at $t+1$ and onwards. The tacit collusion incentive compatibility constraint (TCIC) is therefore

$$\bar{D}_t^* - D_t^* \leq \mathbb{E}_t\left[V_{t+1}^*(\Gamma_{t+1}) - \tilde{V}_{t+1}(\Gamma_{t+1})\right], \tag{13}$$

that is, current gains from defection do not exceed the future loss of value from retaliation. Because of the equivalence between collusive prices and capacity utilization, firms' equilibrium value function is recursively defined by the constrained maximization problem

$$V_t^*(\Gamma_t) = \max_{u_t \in [0,1], I_t} D_t + \mathbb{E}_t\left[V_{t+1}^*(\Gamma_{t+1})\right], \text{ s.t., (13).} \tag{14}$$

To see the effects of the TCIC, note that the immediate cash flow gain from defection is $\bar{D}_t^* - D_t^* = \bar{\Pi}_t^* - \Pi_t^*$ and hence the TCIC (13) becomes

$$(p_t^* - h)A_t(K_t^*)^{\psi}(1 - (u_t^*)^{\psi}) \leq \mathbb{E}_t\left[V_{t+1}^*(\Gamma_{t+1}) - \tilde{V}_{t+1}(\Gamma_{t+1})\right]. \tag{15}$$

(15) implies that gains from tacit collusion require excess capacity, that is, $u_t^* < 1$, else $V_{t+1}^* = \tilde{V}_{t+1}$.

Letting ς_t be the Lagrange multiplier for the TCIC, the optimality conditions with respect to

[8] The situation is more subtle if there are capacity depreciation costs of higher capacity utilization, a case we will consider in Section 5.2 below.

(u_t^*, I_t^*), respectively, are (see Online appendix C.2)

$$\frac{\partial D_t}{\partial u_t} = \frac{\varsigma_t}{(1+\varsigma_t)} \frac{\partial \bar{D}_t}{\partial u_t}, \tag{16}$$

$$-\left[1+\varphi\left(\frac{I_t}{K_t}-\delta\right)\right] + \mathbb{E}_t\left[\frac{\partial V_{t+1}^*}{\partial I_t}\right] = \frac{\varsigma_t}{(1+\varsigma_t)}\left\{-\left[1+\varphi\left(\frac{I_t}{K_t}-\delta\right)\right] + \mathbb{E}_t\left[\frac{\partial \tilde{V}_{t+1}}{\partial I_t}\right]\right\}. \tag{17}$$

Equations (16)-(17) clarify the effects of TCIC on the equilibrium. In the absence of the TCIC or when it not binding ($\varsigma_t = 0$), we obtain the unconstrained optimum for capacity utilization (u_t) and the standard Euler condition for investment (I_t). But with a binding TCIC, higher capacity utilization ceteris paribus lowers current dividends from defection, $\bar{D}_t$, by reducing the industry price p_t and hence the gains from defection to full capacity (see (15)). The optimal u_t in (16) thus takes into account its effect on relaxing the TCIC. Next, eliminating the Lagrange multiplier from (16)-(17) and rearranging terms yields the Euler condition

$$\left[1+\varphi\left(\frac{I_t}{K_t}-\delta\right)\right](1-\chi_t) = \mathbb{E}_t\left[\frac{\partial V_{t+1}^*}{\partial I_t}\right] - \mathbb{E}_t\left[\frac{\partial \tilde{V}_{t+1}}{\partial I_t}\right]\chi_t, \tag{18}$$

where $\chi_t \equiv \left(\frac{\partial D_t}{\partial u_t}\right) \big/ \left(\frac{\partial \bar{D}_t}{\partial u_t}\right) = \left(\frac{u_t^{\psi}}{p_t}\right)(p_t - \sigma(p_t - h))$ and represents the shadow value of using investment to relax the TCIC. Now, investment impacts defection incentives through its effects on (defection) dividends $\bar{D}_t$ and on the punishment (BN) equity value next period because $\tilde{V}_{t+1}$ depends on firms' capacity, K_{t+1}. Equation (18) therefore quantifies the effects of tacit collusion on optimal investment. It also follows that the equilibrium path real decisions (u_t^*, I_t^*) cannot be determined independent of the equilibrium *and* OEP equity values (S_t^* and $\tilde{S}_t$) because of the TCIC.

We note that (18) suggests novel effects of real frictions—specifically, fixed and capital adjustment costs—on equilibrium policies through the tacit collusion channel. To explicate, suppose there is defection at t, so that firms utilize the BN-equilibrium investment policies $(\tilde{I}_\tau)_{\tau \geq t+1}$. But these policies and, hence equity values $\tilde{V}_\tau$, are affected by fixed and adjustment costs (see (12)). For example, because adjustment costs are strictly convex in I/K (see (9)), they will be affected by capacity K (Abel 1981); hence, investment along the equilibrium path will impact adjustment costs in the punishment equilibrium. Thus, in addition to the usual channels, real frictions influence equilibrium investment through their effects on off-equilibrium behavior. More formally,[9]

$$\mathbb{E}_t\left[\frac{\partial V_{t+1}^*}{\partial I_t}\right] = \mathbb{E}_t\Bigg[\Omega_{t,t+1}\Bigg\{\frac{\partial \Pi_{t+1}^*}{\partial K_{t+1}} - m + 0.5\varphi\left(\left(\frac{I_{t+1}^*}{K_{t+1}}\right)^2 - \delta^2\right) + (1-\delta)\left(1+\varphi\left(\left[\frac{I_{t+1}^*(1-\chi_{t+1}) + \chi_{t+1}\tilde{I}_{t+1})}{K_{t+1}}\right] - \delta\right)\right)\Bigg\}\Bigg]. \tag{19}$$

Comparing (18)-(19) and the BN-equilibrium Euler equation (12) above helps clarify the "non-standard" effects of real frictions on investment in the tacit collusion equilibrium: The "standard effects" are captured by terms in the first row of (19) while the effects due to tacit collusion are reflected in the presence of OEP investment $\tilde{I}_{t+1}$ and the TCIC-related terms χ_t and χ_{t+1}.

The tacit collusion channel suggests that firms' risk exposure due to fixed and adjustment costs, which is highlighted in the literature (Jermann 1998; Carlson et al. 2004), will also be affected by the interaction of these real frictions with off-equilibrium-path behavior. For example, higher adjustment costs will ceteris paribus impede firms' capacity modification to address the greater price competition in the BN equilibrium. In a similar vein, higher fixed costs ceteris paribus raise value exposure to lower profits of the BN equilibrium. We will build on this intuition below to examine aggregate effects of real frictions and to derive predictions on the relation of industry concentration and returns.

Now, the TCIC will be binding as long as the monopoly outcome is not incentive compatible at the firm level. However, in our model with a continuous state space, the monopoly outcome will not generically be incentive compatible for realistic number of industry firms even in concentrated oligopolies. In the other polar case, it can be shown—consistent with intuition—that $u_t^* < 1$ is not sustainable when the number of industry firms gets sufficiently large (Online appendix C.3). Thus, the TCIC condition will generically apply for oligopolistic sectors except for low concentration levels; our analysis will henceforth focus on equilibrium paths with binding TCIC.

In sum, the equilibrium path real and financial outcomes (u_t^*, I_t^*, S_t^*) are determined from three conditions: the Euler equation (18); the (binding) TCIC (13); and the asset market clearing condition $\mathbb{E}_t\left[\Omega_{t,t+1}R_{t+1}^*\right] = 1$, where $R_{t+1}^* = (D_{t+1}^* + S_{t+1}^*)/S_t^*$. But, as noted by Cochrane (1991) and Liu et al. (2009), equilibrium returns can also be computed through the returns on investment (ROI) because

[9] Detailed characterization of the equilibrium path is given in Online appendix C.2. In particular, the detailed representation of the Euler condition (18) is given in Equation (C2.27).

the firm's optimality conditions imply $\mathbb{E}_t\left[\Omega_{t,t+1} ROI^*_{t+1}\right] = 1$. In particular, using (18),

$$ROI^*_{t+1} = \frac{\frac{\partial V^*_{t+1}}{\partial I_t} - \frac{\partial \tilde{V}_{t+1}}{\partial I_t}\chi_t}{\left[1 + \varphi\left(\frac{I_t}{K_t} - \delta\right)\right](1 - \chi_t)}. \tag{20}$$

4 Model dynamics and outcomes

We now quantitatively analyze the equilibrium responses of salient real and financial model variables to the fundamental productivity shocks $\{A_{jt}\}_{j,t}$,[10] utilizing perturbation methods (Jin and Judd 2002). Specifically, we undertake quadratic approximations around the deterministic steady state of the model,[11] an approach widely adopted by the literature (Schmitt-Grohe and Uribe 2004; Kim et al. 2008); and desirable properties of the accuracy of such approximations are documented in the literature (Caldera et al., 2011; Fernández-Villaverde et al. 2016; Andreasen et al. 2017).

Similar to the literature (Baxter and Curcini 1993; Boldrin et al. 2001), we implement our quantitative model in a two-sector setting. To focus on the role of tacit collusion, we consider a concentrated sector (sector 1) with tacit collusion and a low-concentration sector (sector 2) where tacit collusion is not sustainable and there is Bertrand-Nash price competition. The BN equilibrium path behavior not only serves as a useful benchmark for equilibrium outcomes in the concentrated sector but also provides intuition on the effects of productivity shocks on its off-equilibrium behavior.[12]

4.1 Calibration

We take industry-level data on capital, investment, output and productivity from the NBER-CES manufacturing database. The latest data available are for 1958-2018 (annual). However, because productivity data are generally available only through 2016, our sample period is 1958-2016. Consistent

[10] For parsimony, our framework follows the multisector general equilibrium literature (e.g., Horvath 2000) in focusing on sectoral productivity shocks as the primary sources of risk in the economy. Empirically, sectoral productivity shocks are a major component of aggregate shocks and the importance of the former relative to the latter appears to be increasing in recent decades (Foerster et al. 2011). Our approach is also consistent with Gabaix (2011) who emphasizes the microfoundations of aggregate shocks.

[11] As mentioned above, (9) implies that there are no adjustment costs in the deterministic equilibrium and OEP steady states. And, consistent with the received RBC literature, we assume that there are no other adjustment costs in the equilibrium steady state. We provide details on the OEP and equilibrium steady states in Section C.4 of the Online appendix.

[12] For computational tractability, we take a linear approximation of the (OEP) Bertrand-Nash general equilibrium—when the concentrated sector also follows the BN equilibrium path—around its deterministic steady state. This analysis yields the policy functions for investment ($\tilde{I}_{1t}$) and ex-dividend equity price ($\tilde{S}_{1t}$) that allow us to compute the (binding) TCIC (13) for the equilibrium path simulations through quadratic approximations around the steady state when there is tacit collusion in the concentrated sector. We utilize Dynare software (version 5.4) for our computations.

with our framework, we focus on consumer goods industries. We obtain times-series data on capacity utilization in U.S. manufacturing industries and consumer goods industry production from the Federal Reserve Board (FRB). We extract information on aggregate variables that match quantities in our model from National Income and Product Accounts (NIPA). Specifically, we use real per capita personal consumption expenditures on durable and non-durable goods to quantify C, and (nominal) per capita personal income on assets (or "financial income")—that is, dividends and interest income—to calibrate W (see (2)). We obtain population data from the U.S. Census Bureau and the price deflator for personal consumption expenditures on goods (P) from the Bureau of Economic Affairs (BEA). In addition, we obtain data on returns from Kenneth French's website.

Table 1 summarizes the calibration for the baseline model. In the standard fashion, certain parameters are externally calibrated, while others are internally calibrated to help match equilibrium steady state outcomes and the model's simulated moments with the data. We use a 1% annual discount rate, that is, $\alpha = 0.99$, which yields a steady state annual real risk-free rate of 1%, which is close to the average real risk-free rate of 0.91% during 1960-2016. There is no consensus on parameterization of relative risk aversion and the IES in the literature. We use a risk aversion of 5, which is similar to Jermann (1998), and is close to the midpoints of the range of risk aversion (2-10) considered reasonable by Mehra and Prescott (1985), as well as the range of risk aversion estimates (2-7) in the finance literature (Elminejad et al. 2022). We set the IES (η^{-1}) at 1.9, which is consistent with the parameterization in recent production-based asset pricing models (e.g., Croce 2014). We internally calibrate the intratemporal ES $\sigma = 5.99$, which is in the range of ES estimates in Redding and Weinstein (2020) and Broda and Weinstein (2006).

We construct the concentrated (or tacit collusive) sector with manufacturing industries that are in the top quintile of sales concentration, defined as the industry sales share of the largest eight firms; correspondingly, the low-concentration sector (or BN equilibrium) is constructed by industries that are below the median sales concentration. The annual depreciation rates in the two sectors are in the $5\%-5.3\%$ range, consistent with their mean *investment rate* (IR), or the investment-to-capital ratio (I/K), in the steady state, which equals the depreciation rate. This calibration is also consistent with the RBC literature (Gomme and Rupert 2007). We calibrate the autocorrelation coefficients of sectoral productivity shocks ρ_j (see (7)) from the data. Meanwhile, there is a wide variation in the literature regarding estimates of the capital adjustment cost parameter (φ), ranging from low to large values over 20 (see Cooper and Haltiwanger 2006). In particular, Liu et al. (2009) report significant estimates of 7.7 and 1 using different estimation approaches. We therefore calibrate adjustment costs in the concentrated sector at 8 and the low-concentration sector at 1.[13] The production elasticities of capital ($\psi_j, j = 1, 2$) in the two sectors (0.49 and 0.46, respectively) are consistent with strictly decreasing returns; these parameters, along with marginal and fixed costs, are internally calibrated.

The variance-covariance matrix of the productivity shocks is internally calibrated to match the volatility of real per capita consumption growth for the baseline model. The assumed annual volatilities of productivity shocks in both sectors are conservative; that is, they are lower than the range utilized in the RBC literature (Gomme and Lakhagvasuren 2012) as well as in the aggregate capital investment literature (Bachman et al. 2013).[14] The negative correlation between the sectoral productivity shocks is consistent with the data and with the conservative approach adopted here.

4.2 Productivity shocks and equilibrium responses

To develop intuition on the effects of tacit collusion on model dynamics, we utilize impulse response functions (IRFs) and examine equilibrium path responses of salient firm and industry level variables to (one standard deviation) sectoral productivity shocks, which are displayed below in Figures 1-3. We follow these responses over a ten year horizon (horizontal axis); and, in the standard fashion, the vertical axis represents the displacement from the "pre-shock" equilibrium path. These responses also determine, through the general equilibrium mechanism, the endogenous responses of aggregate variables in the model, namely, (W_t, P_t, C_t) as well as the risk-free rate $(r_{f,t+1})$.

Figure 1 displays the dynamics of equilibrium responses to a positive productivity shock in the concentrated sector (ε_{1t}^{+}). We note that ε_{1t}^{+} has a positive effect on equilibrium equity price S_{1t} greater than its impact on the OEP equity price $\tilde{S}_{1t}$. Consequently, the TCIC is relaxed (see (15)),[15] and the equilibrium capacity utilization (u_{1t}) falls toward the monopoly level, consistent with intuition on u_1 and tacit collusion provided in Section 3. This raises the industry price (p_{1t}) and sales $(p_{1t}Y_{1t})$, thereby increasing dividends D_{1t}, other things being equal. Higher dividends raise aggregate financial income (W_t), which has a reinforcing effect on industry price (see (11)).

[13] We use a lower value of φ for the non-concentrated sector to be consistent with economic survival in an industry with implicit higher rates of entry.

[14] Recall that the relationship between quarterly (q) and yearly (y) volatility is $\lambda_y = \lambda_q\sqrt{(1+\rho+\rho^2+\rho^3)}$. Using aggregate data, Gommes and Lakhagvasuren (2012) estimate a quarterly productivity shock volatility of 0.86% with $\rho = 0.96$, which implies an annual volatility of 1.66%. Meanwhile, Bachman et al. (2013) utilize SIC-3 manufacturing data to calibrate a quarterly volatility of 2.73% with $\rho = 0.86$, which implies an annual volatility of 4.91%.

[15] Note that the TCIC (13) is defined in terms of levels of on- and off-equilibrium equity values. Hence, the displacements of these equity values from the "pre-shock" equilibrium path quantify the effects of the shock on the TCIC.

In a related vein, there is a mutually reinforcing relation of p_{1t} and the aggregate price index P_t. Indeed, because of the positive effects of W_t and P_t on industry prices, we also find a positive general equilibrium effect of ε_{1t}^{+} on the product price of the low-concentration sector (p_{2t}) that follows a similar trajectory as p_{1t}. This reinforces the observed positive response of P_t to ε_{1t}^{+}. Now, because $C_t = \frac{W_t}{P_t}$, the net effect of ε_{1t}^{+} on consumption is theoretically ambiguous. We find, however, that consumption responds immediately and positively to ε_{1t}^{+}.

The effects of ε_{1t}^{+} on the variables annunciated above are long-lasting even though they begin a recovery to the original pre-shock levels after the immediate response. In particular, because of the reversion process, the expected growth rate of consumption $\left(\mathbb{E}_t\left[\ln\left(\frac{C_{t+1}}{C_t}\right)\right]\right)$ falls, which raises the expected SDF, other being equal. This has implications for other variables of interest.

First, ε_{1t}^{+} raises the investment rate (IR) because of a positive *efficiency* effect. Because productivity shocks are persistent, higher current productivity ceteris paribus increases expected marginal productivity of investment. Moreover, expected marginal value of investment rises because of the increased expected SDF. This effect of ε_{1t}^{+} on IR is also long-term. Indeed, while investment is ceteris paribus negatively related to dividends, in the situation at hand ε_{1t}^{+} raises D_{1t} because of its strong positive effect on sales. Second, the real risk-free rate ($r_{f,t+1}^{r}$) falls. Third, higher expected SDF and dividends are consistent with the positive effect on stock price (S_{1t}), which is not surprising because the firm-, industry- and aggregate-level variables are jointly determined in equilibrium. And, fourth, because of higher stock price, the expected equity return falls; indeed, even the risk premium falls, despite the lower real risk-free rate.

Next, Figure 2 shows the general equilibrium effects of positive productivity shocks (holding other things fixed) in the more competitive, *low-concentration* sector (ε_{2t}^{+}) on tacit collusion in the concentrated sector. In the BN equilibrium, there is an immediate negative output effect of ε_{2t}^{+} on industry price (p_{2t}) because firms produce at full capacity. Consequently, because of the intratemporal substitutability between products of the two sectors, there is a downward effect on the product price (p_{1t}) and hence the aggregate price index P_t falls. Because of productivity persistence, these negative price effects are long term. Hence, there is an adverse efficiency effect on the concentrated sector investment (IR_1). But because of significantly lower sales ($p_{1t}Y_{1t}$), dividends are reduced resulting in lower aggregate income W_t and consumption C_t. Again, due to persistence, the expected SDF falls, resulting in lower stock prices in both sectors (S_{1t}, S_{2t}), as well as off-equilibrium path stock price $\tilde{S}_{1t}$. But in contrast to the earlier analysis, the TCIC now tightens resulting in higher capacity utilization.

Moreover, in response to the lower expected SDF, the real risk-free rate rises, but the concentrated sector risk premium still increases because of the significantly lower S_{1t}. Finally, Figure 2 helps clarify the effects of productivity shocks when the *concentrated sector* is in the OEP punishment (BN) equilibrium and hence provides intuition on the effect of ε_{1t}^{+} on the OEP equity price $\tilde{S}_{1t}$ observed in Figure 1.

In sum, the IRF analysis above helps clarify the effects of tacit collusion and capital on firms' and financial markets' equilibrium response to economy-wide fundamental shocks. Consistent with the representation of defection incentives in the TCIC (see (13)), both on- and off-equilibrium responses of equity markets to shocks and production capacities are central channels for firm- and aggregate-level real effects of shocks. Thus, firms' risk exposure to shocks is materially amplified through their (shocks') effects on defection incentives. Our analysis below will quantify this intuition.

4.3 Moments and correlations of simulated variables

We now present the moments and correlations of simulations from perturbations around the steady state.[16] We also provide the corresponding figures from the data. Panel A of Table 2 displays the simulated volatilities of capacity utilization (u_1) as well as log changes in investment rates in the concentrated sector ($g_{ir,1}$) and the low-concentration sector ($g_{ir,2}$). For the endogenous aggregate variables, we report the volatilities of log changes in aggregate real consumption, financial income and price index (g_c, g_w, g_π).

With tacit collusion, the volatility of capacity utilization generated by the model is 9.35%, compared with the 18.11% in the data.[17] However, the volatility of log changes in the investment rate (g_{ir1}) is 8.9%, which is lower than the data (15.7%). But despite significantly higher volatility of productivity shocks, lower adjustment costs, and similar production elasticity parameterization, the investment rate volatility in the low-concentration sector (g_{ir2})—that follows the BN equilibrium path—is far lower than in the data. And, of course, the volatility of capacity utilization is counterfactually zero in this sector. In addition, untabulated results show that the mean investment-to-sales

[16] The steady state equilibrium capacity utilization in the concentratrated sector (54%) implies that the average capacity utilization in our two-sector economy is 77%, which is close to the mean 79.6% capacity utilization in the U.S. manufacturing sector in our sample period. We note that the FRB data on capacity utilization only cover a relatively small number of industries and there is sparse representation of the industries that comprise our concentrated sector. We therefore use the capacity utilization in the manufacturing sector as the benchmark. And as mentioned above, through the choice of the depreciation rates, the steady state investment rates (IR) in both sectors closely match their respective sample means.

[17] This is the annualized volatility of the raw monthly manufacturing sector capacity utilization data from FRB during 1962-2016.

ratio (ISR) in the concentrated sector is 0.01 compared with 0.03 in the data; in contrast, the mean ISR in the low-concentration sector is lower than the data by a factor of 100. We conclude that the tacit-collusive equilibrium path fits capacity utilization and investment data significantly better than the BN equilibrium path. For parsimony, we will henceforth focus on the concentrated sector.

For the aggregate variables, the volatility of log changes in real consumption generated by the model (2.7%) is close to the data (2.6%). But the baseline model understates the volatility of log changes in financial income (3.4%) and aggregate price index (1%) relative to the data (6.3% and 2.8%, respectively).

In Panel B, we display the correlation of firm-level and aggregate variables (u_1,IR_1, W, P, C) generated by the model. The sectoral investment rate (IR_1) is positively correlated with (W, P, C), which is similar to the data, but the model correlation coefficients are significantly lower than in the data. Equilibrium capacity utilization in the model is negatively correlated with (W, P, C) and this is similar to the data, although the model correlation coefficients are significantly higher than in the data.[18] In untabultaed results, we also find that, similar to the data, the model generates a very high correlations among W and P and C.

Panel C analyzes the four-year autocorrelation coefficients (ACF) of log changes in salient concentrated sector and aggregate variables. In untabulated results, the model matches the high positive autocorrelation coefficients of capacity utilization (u_1) and the aggregate variables (W, P, C) in the data. Turning to Panel C, the model generates negative serial correlation in investment rates (g_{ir1}), similar to the data (except for year $t-3$). There is negative but very low autocorrelation in consumption (g_c), which is distinct from the positive $ACF(1)$ and $ACF(2)$ in the data, but similar to the negative and low $ACF(3)$ and $ACF(4)$ in the data. The autocorrelations in financial income (g_w) and untabulated price index (g_π) in the baseline model are not consistent with the data. As we will show below, however, the autocorrelation of g_w generated by the extended model with entry and exit is closer to the data.

Panel A of Table 3 displays simulated model moments and correlation coefficients of the concentrated sector returns and the economy-wide real risk-free rate. The mean real risk-free rate is 0.86%, which is very close to the 0.91% average real risk-free rate in our sample, in contrast to the well known "risk-free rate puzzle" (Weil 1989). The volatility of the risk-free rate (0.11%) is significantly

[18] The literature documents pro-cyclical capacity utilization by using national income, and not per capita *financial* income. To our knowledge, the literature has not investigated correlation of manufacturing capacity utlization with the price index and per capita consumption of durables and non-durables.

lower than the corresponding volatility in the data (2.15%), however. The mean equity risk premium (ERP) is 2.87% versus the mean value-weighted industry portfolio ERP of 6.94% in our sample. The model generates ERP volatility of 11.87% compared with 22.95% in the data. As benchmarks from the macrofinance production-based asset pricing literature, Croce (2014) uses a risk aversion of 10 and IES of 2 and finds significantly higher risk-free rate and lower ERP without long-run productivity risks (the comparable setting to our model). Jermann (1998) utilizes additive utility preferences and finds a significantly higher risk-free rate and lower ERP with adjustment costs but no habit formation. Thus, strategic interaction among firms appears to induce additional volatility in the SDF given that the volatilities of consumption and income generated by our model are either close to or understate the data (Table 2).

In Panel B of Table 3, we show the autocorrelations of the real risk-free rate and the ERP. The model generates positively autocorrelated risk-free rate, which is similar to the data. However, the ERP generated by the model also is positively correlated; in the data, $ACF(1)$ and $ACF(2)$ are negative while $ACF(3)$ and $AFC(4)$ are positive.

Overall, the baseline quantitative analysis shows that the effects of fundamental shocks on defection incentives can raise mean firm-level equity risk premia (ERP) and lower aggregate real risk-free rate (compared with benchmark models) even when consumption volatility matches the data. The baseline model understates the volatilities of firm-level capital investment and ERP, as well understating volatilities of aggregate income, price index and real risk-free rate. We now turn to examine the effects of entry and exit.

4.4 Entry and exit

We now consider the interaction of entry and exit with tacit collusion by analyzing model dynamics when there is entry and exit. Endogenous entry along the tacit collusion equilibrium path occurs because of productivity shocks that potentially raise expected industry profits. Given Γ_t, a finite pool of symmetric firms $(1, ..., N_1^e)$ is sequentially given the opportunity to enter after expending sunk entry costs (ζ_1) and investing in production capacity that becomes available at $t + 1$, consistent with the timing conventions of the model and with the literature (Ericson and Pakes 1995). Similar to the literature, in our quantitative analysis, we focus on symmetric equilibrium paths (Bilbiie et al. 2012). In particular, entry capacity size is symmetric with incumbents as in Brock (1972) and Smith (1974), and firms enter at t with capacity choice $K^*_{1,t+1}$. In practice, there is displacement of less efficient firms (possibly due to idiosyncratic shocks) by entrants and, indeed, the asset pricing literature highlights displacement risk (Berk et al. 2004; Garleanu et al. 2012). Similar to the literature (Bilbiie et al. 2012; Corhay et al. 2020), we model displacement risk along the tacit collusion equilibrium path through an exit rate $x_1 \in (0,1)$. The "law of motion" of the number of firms is $N_{1,t+1} = (1-x_1)(N_{1t} + e_{1t})$, where e_{1t} is the number of entrants at t. In equilibrium, entry occurs as long as $(1-x_1)\mathbb{E}_t\left[V^*_{1,t+1}\right] - (K^*_{1,t+1} + \zeta_1) \geq 0$.

We choose entry cost calibration in the concentrated sector to allow a meaningful quantitative comparison of equilibrium path outcomes with those of the baseline (no entry) model considered above;[19] this is useful to build intuition on the effects of entry and exit with tacit collusion and capital investment. The exit rate is internally calibrated at 1% to reduce the distance between the second moments and ACF generated by the model with the data. Apart from the exit rate and entry cost, the model calibration is the same as in Table 1.

Because of exit, industry capacity falls (unless offset by sufficient entry), which ceteris paribus raises product prices and hence expected investment returns (ROI_{t+1}). This increases optimal investment and hence capacity sizes of *incumbent* firms, other things being equal. Thus, exit has conflicting effects on entry in a production economy with endogenous capacity: On the one hand, exit induces entry by raising expected returns of incumbent firms; on the other hand, it deters entry because the required initial capacity ($K_{1,t+1}$) increases.[20]

Figure 3 shows the equilibrium response of some salient variables to a positive productivity shock in the concentrated sector (ε^+_{1t}). We start by focusing on the effects on entry. Because of the positive efficiency effect (noted earlier), investment rate (IR_1) increases and hence K_{t+1} rises. The latter raises entry capacity costs and the initial effect of ε^+_{1t} on entry is negative. But this reduces the number of incumbents in the next period, raising expected returns and entry rate subsequently rises so that the number of entrants (e_1) reverts to the initial levels after two years. This "reverting" effect of productivity shocks on entry rate is an implication of endogenous capacity entry costs. But because of exit, the number of incumbents is still below the pre-shock levels, which raises product prices in subsequent periods allowing capacity utilization (u_1) to fall (as explained previously). As the number

[19] We keep the low-profit low-concentration sector and the off-equilibriun path (OEP) "closed" with a fixed number of firms, as in the baseline model considered earlier. The entry cost is calibrated so that entry is not optimal with the same number of firms and steady state outcomes as in the baseline case. These assumptions ensure that we characterize the equilibriium path with entry and exit through perturbations around the same steady and OEP policies as in baseline model. We find, however, that allowing OEP entry and exit adds to computational complexity but does not significantly affect the outcomes.

[20] But these effects of exit are not present, however, in entry models without production capacity costs of entry, such as Bilbiie et al. (2012), Corhay et al.(2020) and Loualiche (2021).

of incumbents recovers due to continued entry, the positive effect of ε_{1t}^{+} on p_1 reaches its maximum about two years after the shock and then begins a slow decline. Consequently, the negative effect of ε_{1t}^{+} on u_1 also reaches its maximum after two years.

Meanwhile, because of higher output, there is an immediate downward effect of ε_{1t}^{+} on the current industry price (p_{1t}), which recovers quickly as in the subsequent period because the number of incumbents is still below the pre-shock level. Due to intratemporal product substitutability, however, the lower p_{1t} reduces demand for the sector 2 product, leading to a significant reduction in p_{2t} and hence D_{2t}. Consequently, the aggregate implications of ε_{1t}^{+} are that P_t and W_t fall (despite the rise in dividends of the concentrated sector) but revert to their original values two years after the shock, in line with the price and entry trajectories. C_t also falls but begins recovery only after two years so that the expected SDF rises due to lower expected growth of consumption. Hence, the real risk-free rate initially falls but then recovers at $t+2$ as consumption growth reverses the initial effect. In contrast, ε_{1t}^{+} has a strong positive effect on equity price S_{1t} because of higher expected SDF and dividends. Because of the higher S_{1t}, the ERP falls—despite the lower risk-free rate—but recovers by the second year after the shock.

A visual comparison of the trajectories of the salient model variables in Figures 1 and Figure 3 indicates greater volatility in the presence of entry and exit, as the post-shock trajectories get reversed relatively sharply. This intuition is verified in the simulation outcomes with entry and exit reported in Table 4. For expositional parsimony, we display results of variables that are most significantly different from the baseline model outcomes (Tables 2 and 3). Panel A of Table 4 shows that, relative to the baseline results, entry and exit significantly raise the volatilities of concentrated sector capacity utilization, investment rates, as well as consumption; indeed, these now exceed those in the data. The volatility of aggregate price index and financial income also rise, with the former essentially matching the data, while the volatilities of investment rates and financial income are closer to the data relative to the baseline model.

Panel B of Table 4 shows significant changes in the autocorrelation patterns of log changes in consumption and income (relative to the baseline). Log changes in consumption are now significantly positively autocorrelated for years $t-1$ and $t-2$, similar to the data. The positive autcorrelation over longer horizon (years $t-3$ and $t-4$) does not match the data, however. A major change (from the baseline) is the autocorrelation pattern of log changes in financial income: These are now significantly positively autocorrelated, similar to the data. Untabulated analysis shows, however, that the ACFs of the aggregate price index are still not consistent with the data and the correlation of the investment rate (IR_1) with W and P is weakly negative (with entry and exit).

There is an intuition that with higher consumption and, hence, value volatility, endogenous entry in the presence of tacit collusion and capital will amplify firms' risk exposure to fundamental shocks. In Panel C, we indeed find that entry and exit improve the empirical performance of financial returns compared to the baseline model above. The mean sectoral equity premium is 5.2%, which is much closer to the data (6.9%) compared with baseline model and the volatility of ERP is also substantially higher. Moreover, the risk-free rate is still close to the data (0.94% versus 0.91%) and its volatility is now much closer to the data.

We conclude that entry and exit improve the fit of the model with data along some dimension, but not so in others, consistent with the view that the data reflect a large number of industries with widely varying rates of entry and exit. Our analysis highlights the effects of variations in oligopolistic industry structure due to endogenous entry and idiosyncratic exit on aggregate dynamics.

4.5 Micro and aggregate effects of adjustment costs

In Section 3.1, we noted the effects of real frictions on equilibrium outcomes through the tacit collusion channel. We now quantitatively analyze firm- and aggregate-level effects of variations in capital adjustment costs. Specifically, we simulate our economy with entry and exit (Section 4.4) when capital adjustment costs are significantly higher in the concentrated sector relative to the baseline calibration (φ_1 is raised to 14 from 8, holding other parameters the same). We display simulated moments of salient variables in the second column of Table 5, and use Table 4 for comparison. Not surprisingly, volatility of log changes in investment rate falls significantly and the volatility of capacity utilization is also reduced. We find, however, that entry becomes more lumpy with higher capacity adjustment costs. Untabulated results show that mean entry per period falls while its volatility rises and, consequently, the volatilities of log changes in industry product price and dividends increase. We thus see higher volatility of log changes in W, P and C. In addition, because of the increased friction in accommodating productivity shocks both on- and off-equilibrium path, the mean and volatility of firm-level risk premium as well as the real risk-free rate rise. In sum, real frictions further amplify the effects of endogenous entry on firms' risk exposure along the tacit collusion equilibrium path.

A useful benchmark for these results is the influential literature that examines the aggregate effects of firm- or plant-level real frictions (Veracierto 2002; Khan and Thomas 2008). This literature generally finds that in competitive settings such frictions are essentially irrelevant for aggregate dynamics in general equilibrium because prices adjust to compensate for the frictions. In particular, Khan and Thomas (2008) analyze a model with non-convex adjustment costs in a competitive setting and do not consider entry with capacity and sunk fixed costs. They conclude that adjustment costs are essentially irrelevant for aggregate dynamics once factor supplies are allowed to adjust in general equilibrium. Our model assumes infinitely elastic supply of investment goods and smooth adjustment costs. Nevertheless, we find significant effects of firm-level adjustment costs on aggregate dynamics with tacit-collusive prices and endogenous entry with fixed and capacity costs. Intuitively, the compensating price mechanisms that appear to be responsible for the neutrality results in the literature are restricted due to tacit collusion and real frictions in entry.

4.6 Micro and aggregate effects of productivity processes

We examine next the effects of volatility and persistence of productivity shocks through the tacit collusion channel. Column 3 of Table 5 displays the results when the productivity persistence of the concentrated sector is varied from 0.73 (in the baseline) to 0.63, holding other parameters the same; and Column 4 repeats the same procedure when productivity volatility in this sector is reduced to 0.05%. We find that variations in the parameters of the productivity process in the concentrated sector have marginal effects. In contrast, as seen in Columns 4 and 5, relatively small variations in the productivity persistence (from 0.96 to 0.95) and volatility (from 1.5% to 1%) in the low-concentration sector have substantial effects on the tacit collusion equilibrium of the concentrated sector, especially on the risk premium. In Section 4.2 above, we qualitatively highlighted the cross-sectoral or general equilibrium effects of productivity shocks in the low-concentration sector on response functions in the concentrated sector, as well as the aggregate effects of these shocks. The analysis in Table 5 indicates that these general equilibrium effects are quantitatively quite significant.

5 Industry concentration and equity returns

As mentioned before, the literature presents conflicting evidence on the cross-sectional relation of industry concentration and equity returns.[21] In this and the next section, we show—conceptually, quantitatively and empirically—that tacit collusion with endogenous capital can help resolve the observed ambiguity in the concentation-return relation. Now, the cross-sectional regressions of returns on industry concentration are partial, or industry, equilibrium in that they take as *given* concentration

[21] We are grateful to an anonymous referee for highlighting this issue to us.

and other covariates that affect returns. That is, these regressions empirically examine comparative statics on returns with respect to exogenous variations in variables that are (at least temporarily) taken as given. We therefore analyze the implications of tacit collusion and capital investment on the concentration-return relation in this setting.[22]

Intuitively, industry production capacity plays a central role in tacit collusion because, as seen above, defection incentives are determined by firms' production capacities both on and off the equilibrium path. In particular, along the equilibrium path, defection incentives are positively related to excess capacity, while these incentives are negatively related to industry capacity in the OEP punishment path. But industry capacity is strongly affected by concentration. Hence, there is a clear link between concentration and risk premium through the effects of the former on defection incentives. More concisely, let us denote, for any variable H_t, $\Delta_N(H_t) \equiv H_t(N+1) - H_t(N)$. Then it can be shown using the TCIC (13) (see Online appendix D) that the effect (in terms of sign) of increasing the number of industry firms on equilibrium industry price is

$$\Delta_N(p_t) \propto \Delta_N \left(\mathbb{E}_t \left[V_{t+1}(\Gamma_{t+1}) - \tilde{V}_{t+1}(\Gamma_{t+1}) \right] \right) - (p_t - h)\Delta_N(\bar{Y}_t - Y_t), \qquad (21)$$

where $\bar{Y}_t$ is the firm's production at full capacity. The first term in RHS of (21) represents the *punishment* effect (PE) of lower concentration and is generally positive, while the second term is the *market dilution* effect (MDE) of lower concentration and is negative. Thus, variations in concentration generally have ambiguous effects on p_t. The effects of industry concentration and firm-level dividends, $\Delta_N(D_{t+1})$, and hence the risk premium, $\Delta_N\left(\mathbb{E}_t\left[R_{t+1} - R_{f,t+1}\right]\right)$, are therefore generally ambiguous.

Now, if $\Delta_N\left(\Omega_{t,t+1}\right)$ and $\Delta_N\left(R_{f,t+1}\right)$ are small because of realistically large number of sectors, then the industry $\Delta_N\left(\mathbb{E}_t\left[R_{t+1} - R_{f,t+1}\right]\right) \simeq -\Delta_N Cov_t\left(\Omega_{t,t+1}, \left(\frac{D_{t+1}+S_{t+1}}{S_t}\right)\right) R_{f,t+1}$ (Online appendix D). But if the MDE is strong relative to the PE, then $\Delta_N\left(\frac{D_{t+1}+S_{t+1}}{S_t}\right) < 0$ and the covariance of the SDF with equity payoffs increases (in magnitude) with N, so that $\Delta_N\left(\mathbb{E}_t\left[R_{t+1} - R_{f,t+1}\right]\right) > 0$, that is, the risk premium is negatively related to concentration. An opposing argument applies when the PE is relatively strong, so that $\Delta_N\left(\frac{D_{t+1}+S_{t+1}}{S_t}\right) > 0$ and hence ceteris paribus $\Delta_N\left(\mathbb{E}_t\left[R_{t+1} - R_{f,t+1}\right]\right) > 0$. We now derive testable empirical hypotheses (or predictions) on the relation of concentration and returns by considering firm-, industry- and macro-level characteristics that help resolve the conflicting MDE and PE of concentration on the risk premium.

[22] Such partial equlibrium analysis is often utilized in the literature as part of general equilibrium analysis (Veracierto 2002; Opp et al. 2014)

5.1 Operating leverage, profitability, and adjustment costs

Consider first the case of firms with high operating leverage-to-operating profitability (OL/OP) ratios. Firms with sufficiently high fixed costs, or high OL/OP ratios, will ceteris paribus have low dividends and equity values in low W_t (or high SDF) states. With smooth consumer preferences and production relationships (see Section 2), $S_t^* = \mathbb{E}_t\left[V_{t+1}^*\right]$ is continuous in W_t.[23] Hence, for every $\epsilon_t^s > 0$, there exists $\epsilon_t^w > 0$ such that $S_t^* < \epsilon_t^s$ whenever $W_t < \epsilon_t^w$ (holding fixed other elements of Γ_t). But due to limited liability, S_t^* and $\tilde{S}_t$ are uniformly bounded below by zero. Hence, for OL/OP sufficiently high and W_t sufficiently low, the PE, $\mathbb{E}_t\left[V_{t+1}^* - \tilde{V}_{t+1}\right]$, can be made arbitrarily close to zero. Consequently, for such firms the MDE must dominate the PE in high SDF states because the markup $(p_t^* - h)$ stays strictly positive along tacit collusive equilibrium paths. The argument above then implies that $\Delta_N\left(\mathbb{E}_t\left[R_{t+1}^* - R_{f,t+1}^*\right]\right) > 0$ for firms with high OL/OP ratios. Similarly, in low W_t (or high SDF) states, firms with high adjustment costs ceteris paribus can not reduce their production capacities, lowering $\mathbb{E}_t\left[V_{t+1}^*\right]$ and weakening the punishment effect. Hence, $\Delta_N\left(\mathbb{E}_t\left[R_{t+1}^* - R_{f,t+1}^*\right]\right) > 0$ for such firms.

An opposing argument applies for firms that have high OP/OL ratios because of low marginal costs relative to fixed costs. In this case, S_t^* stays strictly bounded away from zero even in low income states and hence the PE remains significant in high SDF states. Reversing the argument in the previous paragraph, $\Delta_N\left(\mathbb{E}_t\left[R_{t+1}^* - R_{f,t+1}^*\right]\right) < 0$ for firms with sufficiently high OP/OL ratios.[24] In a similar vein, optimal investment (I_t^*) is ceteris paribus negatively related to fixed and capital adjustment costs (see (18)). High investment firms therefore ceteris paribus reflect relatively low fixed and/or low adjustment costs so that $\Delta_N\left(\mathbb{E}_t\left[R_{t+1}^* - R_{f,t+1}^*\right]\right) < 0$ for such firms.

It is useful to quantitatively demonstrate the opposing relation of equity returns and industry concentration based on the foregoing analysis. To do so, we adapt the baseline general equilibrium quantitative model of the previous section (Sections 4.2 and 4.3) to an *industry* equilibrium where aggregate income and price index (W_t, P_t) are taken as state variables (rather than being endogenous). We consider the relation of returns and concentration over a range of fixed cost, marginal cost, and adjustment costs parameterizations. Hence, for computational tractability, we adopt the loglinear approach of Jermann (1998) and Horvath (2000) who also quantitatively analyze returns (or asset

[23] The value functions V_t^* and $\tilde{V}_t$ will be continuous in the state variables from an application of the Theorem of the Maximum or from the arguments in Blume, Easley and O'Hara (1982).

[24] This prediction applies even in the absence of entry. Hence, it is different from Corhay et al. (2020) where a positive relation of concentration and risk premium arises because of entry induced by higher profits in more concentrated industries. The argument here relates to firms with higher operating profitability at a *given* level of concentration.

pricing) in production economies.[25] In this approach, the logs of firms' state (W_t, P_t, A_t), namely, $\boldsymbol{\mu}_t \equiv (w_t, \pi_t, a_t)^\intercal$ follow an AR(1) process, that is, $\boldsymbol{\mu}_{t+1} = \boldsymbol{\rho}\boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_{t+1}$, where $\boldsymbol{\varepsilon}$ is i.i.d mean zero with variance-covariance matrix $\Sigma = [\Sigma_{ij}]$ and $\boldsymbol{\rho}$ is the square diagonal matrix of auto-correlation coefficients with $\text{diag}(\boldsymbol{\rho}) = (\rho_w, \rho_\pi, \rho_a)$, $0 \leq \boldsymbol{\rho} \leq 1$.

Considering the loglinear representation of returns in our framework provides additional intuition on our model. Using the standard Campbell and Shiller (1988) loglinear approximation, equity returns are given by $r_{t+1} = \varkappa_{d0} + \varkappa_{d1}\ell_{d,t+1} - \ell_{dt} + d_{t+1} - d_t$, where ℓ_{dt} is the log price-dividend ratio $\log(S_t/D_t)$, and $\varkappa_{d0}$ and $\varkappa_{d1}$ are approximating constants that depend on the mean level of ℓ_d . As seen above, because of tacit collusion, firms' equilibrium real policies—and hence dividends—depend on the aggregate and industry state variables, their existing capital stock and on- and off-equilibrium stock prices. Consequently, the equilibrium price-dividend ratio is a (linear) function of the logs of aggregate and industry state variables, as well as the endogenous off-equilibrium stock price (see Online appendix D.2). The latter distinguishes our analysis from the standard loglinear return representation in production-based asset pricing models. Heuristically, the coefficients of ℓ_d capture firm and aggregate responses to sectoral productivity shocks through their impact on defection incentives from tacit collusion that have been analyzed in general equilibrium above.[26]

We parameterize the variance-covariance matrix of shocks to $\boldsymbol{\mu}_t \equiv (w_t, \pi_t, a_t)^\intercal$ (that is, Σ) from NIPA data on financial income and consumer price index, as well as the NBER-CES database for consumer goods industries. We maintain the parameterization relating to the discount factor, risk aversion and IES from the previous section; the taste-factor parameter (ϕ) takes a lower value to reflect a generic industry rather than sectors; and the ES (σ) is again in the range estimated by Broda and Weinstein (2006). The calibration is summarized in Table A.1 of the Online appendix. We simulate conditional equity returns for various concentration levels as well as a range of OL, OP, and adjustment costs. The means of simulated returns are displayed in Figures 4 through 6 through (three-dimensional) "surface maps" plotted against industry concentration and model parameters.

The analysis above predicts that for firms with relatively high OL/OP ratios, the mean returns will be positively (negatively) related to number of industry firms (concentration). And the concentration-risk premium relation will reverse in sign—that is, be positive—for firms with rela-

[25] Jermann (1998) and Horvath (2000) study production economies in a competitive setting. A substantial strand of the RBC and macrofinance literature also utilizes log-linear analytic approximations around a deterministic steady state, including dynamic models with tacit collusion (Rotemberg and Woodford 1992).

[26] Methodologically, equilibrium investment and capacity utilization policies have to be determined simultaneously with the stock price, which requires solving a higher-order polynomial compared with the standard case.

tively high OP/OL ratios. We parameterize high OL/OP ratio firms by choosing high ratios of fixed to marginal costs (that is, m/h) and conversely for high OP/OL ratio firms. Figure 4 considers the case of high OL/OP ratio firms and shows that mean returns are monotonically negatively related to concentration, consistent with the predicted relation, for an open interval of high ratios of fixed and marginal costs. Next, Figure 5 analyzes the case of firms with high OP/OL ratios, or low OL/OP ratio firms, by using significantly lower ratios of fixed and marginal costs. In contrast to Figure 4, but consistent with the theoretical prediction above, mean returns are now monotonically *positively* related to concentration. Finally, Figure 6 analyzes the effects of high adjustment costs on mean returns. In this case, we set the OL/OP ratio to be significantly lower than in Figure 4; nevertheless, for high adjustment costs, there is a negative relation of mean returns and concentration, similar to Figure 4. This analysis confirms that high adjustment costs affect the concentration-returns relation along a channel similar to high operating leverage, as we argued in Section 5.1.

5.2 Endogenous capacity depreciation

The RBC literature examines the implications of endogenous depreciation through capacity utilization, but generally in a competitive framework. However, the trade-off between capacity utilization and depreciation will influence tacit collusion by affecting the relative strengths of PE and MDE.[27] Intuitively, because of depreciation costs of higher capacity utilization, $(\tilde{u}_t, \tilde{K}_t)$ will ceteris paribus be lower, raising prices and equity values in the BN equilibrium and weakening the PE. But defecting firms will not necessarily produce at full capacity, which will weaken the MDE.

To show this, we let $K_{t+1} = (1 - \delta(u_t))K_t + I_t$, where $\delta : [0, 1] \to [0, 1]$ is strictly increasing and convex on $(0, 1)$. Similar to Greenwood et al. (1988), we set $\delta(u) = \frac{1}{\xi}u^{\xi}, \xi > 1$, so that the marginal depreciation cost of capacity utilization $\dot{\delta}(u) = u^{\xi-1}$. Hence (for a fixed $u < 1$), both $\delta(u)$ and $\dot{\delta}(u)$ are strictly decreasing in ξ. We verify that $\bar{u}_t^* < 1$ and $\tilde{u}_t < 1$ for ξ not too high (Online appendix E). Our focus is on the effect of ξ on the risk premium. As noted above, there are two opposing forces at play here. Lower ξ (that is, higher $\dot{\delta}(u)$) ceteris paribus weakens the PE by reducing $(\tilde{u}_t, \tilde{K}_{t+1})$. This is easiest to see in the punishment equilibrium steady state where it can be shown that $\frac{\partial \tilde{K}}{\partial \xi} > 0$ when the marginal cost h is not too large (see Online appendix E). Thus, higher depreciation costs (lower ξ) ceteris paribus weaken the PE. On the other hand, lower ξ dampens defection incentives along the equilibrium path because $\bar{D}^* - D^* = (p^* - h)(\bar{u}^* - u^*)$ and

[27] We are grateful to an anonymous referee for pointing this out to us.

$\frac{\partial \bar{u}^*}{\partial \xi} > 0$ in the steady state. The weaker defection incentives allow higher tacitly collusive prices to be sustained in equilibrium with lower industry concentration and weaken the MDE. The net effects of ξ on equilibrium markups and the risk premium are hence theoretically ambiguous. However, because the PE is (potentially) of infinite duration, we expect that the PE will be weaker than the MDE for sufficiently high $\dot{\delta}(u)$. Thus, the prediction is that for industries with high depreciation costs of capacity utilization, $\Delta_N \left(\mathbb{E}_t \left[R^*_{t+1} - R^*_{f,t+1}\right]\right) > 0$.

5.3 Strategic entry deterrence

A theoretical literature highlights strategic entry deterrence through firms with market power (Spence 1977; Dixit 1980). We derive empirical implications for the concentration-returns relation by obtaining from the equilibrium entry condition (Section 4.4) the conditional probability mass function for the number of potential entrants $\Phi^*_{i,t+1} = \Pr\left(e^*_{t+1} = i \middle| \hat{\Gamma}_t\right), i = 0, 1, ..., N^e$. The corresponding incumbents' value functions are denoted $V^*_{t+1}(\Gamma_{t+1}, e^*_{t+1} = i) \equiv V^*_{i,t+1}(\Gamma_{t+1})$. In general, let $\mathbf{K}^*_{t+1}$ be the profile of incumbents' capacities at $t+1$ and partition the state $\Gamma_{t+1} = (\hat{\Gamma}_{t+1}, \mathbf{K}^*_{t+1})$. If the MDE dominates PE, then for all potential entrants, $V^*_{i,t+1}(\mathbf{K}^*_{t+1}, \hat{\Gamma}_{t+1}) < V^*_{i,t+1}(\breve{\mathbf{K}}^*_{t+1}, \hat{\Gamma}_{t+1})$ if $\mathbf{K}^*_{t+1} > \breve{\mathbf{K}}^*_{t+1}$. Optimal entry thus implies that the distribution $\mathbf{\Phi}^*_{t+1} = (\Phi^*_{0,t+1}, ..., \Phi^*_{N_e,t+1})$ is decreasing in $\mathbf{K}^*_{t+1}$ in the sense of first order stochastic dominance if the MDE dominates. Therefore,

$$\frac{\partial \mathbb{E}_t\left[V^*_{t+1}(\Gamma_{t+1})\right]}{\partial I^*_t} = \sum_{i=0}^{N_e} \Phi^*_{i,t+1} \frac{\partial \mathbb{E}_{t,\Gamma}\left[V^*_{i,t+1}(K^*_{t+1}, \hat{\Gamma}_{t+1})\right]}{\partial K^*_{t+1}} + \sum_{i=0}^{N_e} \frac{\partial \Phi^*_{i,t+1}}{\partial K^*_{t+1}} \mathbb{E}_{t,\Gamma}\left[V^*_{i,t+1}(K^*_{t+1}, \hat{\Gamma}_{t+1})\right], \quad (22)$$

where $\mathbb{E}_{t,\Gamma}[\cdot]$ denotes expectations with respect to exogenous state Γ_{t+1}. Now, if the MDE is dominant, then $V^*_{i,t+1}$ is decreasing in i and $\mathbf{\Phi}^*_{t+1}$ is decreasing in $\mathbf{K}^*_{t+1}$; hence, the second term in (22) is positive when the MDE is strong. In this case, we get $K^{e*}_{t+1} = K^*_t + I^{e*}_t > K^*_{t+1} = K^*_t + I^*_t$, where I^{e*}_t and I^*_t denote investment with and without entry threat, respectively. We conclude that in industries where large incumbent capacities deter entry, higher capital investment indicates a strong MDE and hence the effect of investment on the concentration-returns relation should be stronger.

5.4 Time-varying relation of concentration and returns

For analytic tractability, and similar to the common approach in the RBC and asset pricing literatures, we have assumed that the variance-covariance matrix Σ of innovations to the state variables is stationary. However, a large literature considers time-varying risk premia. In particular, Duffee (2005) finds that the conditional covariance between aggregate stock returns and consumption growth is high in times of high aggregate stock returns. Because aggregate returns are weighted averages of firm-level returns, this finding would imply that in our setting, on average, $-Cov_t\left(\Omega_{t,t+1}, \frac{D^*_{t+1}+S^*_{t+1}}{S^*_t}\Big|\hat{R}_t\right) < -Cov_t\left(\Omega_{t,t+1}, \frac{D^*_{t+1}+S^*_{t+1}}{S^*_t}\Big|R_t\right)$, for aggregate return $\hat{R}_t > R_t$. It then follows that $\left|\Delta_N\left(\mathbb{E}_t\left[R^*_{t+1} - R^*_{f,t+1}\right]\Big|\hat{R}_t\right)\right| > \left|\Delta_N\left(\mathbb{E}_t\left[R^*_{t+1} - R^*_{f,t+1}\right]\Big|R_t\right)\right|$. Hence, $\Delta_N\left(\mathbb{E}_t\left[R^*_{t+1} - R^*_{f,t+1}\right]\Big|\hat{R}_t\right) > \Delta_N\left(\mathbb{E}_t\left[R^*_{t+1} - R^*_{f,t+1}\right]\Big|R_t\right) > 0$ for high OL firms, and this inequality will reverse for high OP and high investment firms (Section 5.1).

Another source of time variation in the concentration-return relation is secular trends in industry fundamentals. Markups and industry concentration in U.S. have increased significantly in recent decades (Farhi and Gourio 2018; Autor et al. 2020; De Loecker et al. 2020), and these trends are consistent with our sample (Table A.2 of Online appendix). The literature relates higher markups to technological changes that have raised fixed costs (Berry et al. 2019), consistent with evidence that higher concentration and markups are due to a few efficient ("superstar") firms with higher fixed-to-marginal cost ratios (Autor et al. 2020). For the concentration-return relation, a secular trend of increasing fixed-to-marginal cost ratios should weaken the PE for high OL firms and hence strengthen the negative effect of OL, while superior efficiency of incumbents implies less reliance on capacity for entry deterrence and hence a weakening effect of investment.

5.5 Summary of empirical hypotheses

We now summarize novel empirically testable hypotheses derived in this Section, which we will test below. The first three hypotheses are derived from extending the arguments in Section 5.1. The first part of these hypotheses predicts that the slope of equity returns with respect to industry concentration is monotone in operating leverage, operating profitability and investment. The second part asserts that the concentration-returns relation becomes positive or negative for sufficiently high values of these firm characteristics.

Hypothesis 1. *Industry concentration is ceteris paribus more negatively related to returns of firms with higher operating leverage and the return-concentration relation is negative for high operating leverage firms.*

Hypothesis 2. *Industry concentration is ceteris paribus more positively related to returns of firms with higher operating profitability and the return-concentration relation is positive for high operating profitability firms.*

Hypothesis 3. *Industry concentration is ceteris paribus more positively related to returns of firms with higher investment and the return-concentration relation is positive for high investment firms.*

Next, Section 5.2 predicts that industry concentration will be negatively related to risk premium when marginal depreciation costs of capacity utilization, $\dot{\delta}(u)$, are high.

Hypothesis 4. *Industry concentration is ceteris paribus more negatively related to returns of firms with higher sensitivity of capital depreciation rates to capacity utilization (that is,* $\dot{\delta}(u)$*).*

The next hypothesis is an implication of using capital investment for entry deterrence (Section 5.3):

Hypothesis 5. *In capital-intensive high entry-threat industries, the positive effect of investment on the concentration-returns relation is ceteris paribus strengthened.*

While Section 5.3 and Hypothesis 5 focus on the effects of entry deterrence in capital-intensive industries, deterrence can also more generally occur through large size of incumbents (Siegfried and Evans 1994; Gershon 2013). In this case, higher OL ceteris paribus increases profits by deterring entry, weakening the relative strength of MDE over PE for high OL firms (see Section 5.1). Thus:

Hypothesis 6. *In high entry-threat industries where large incumbent size deters entry, the negative effect of higher operating leverage on the concentration-returns relation is ceteris paribus weakened.*

Next, considering time-varying risk premia (see Section 5.4) yields the hypotheses that:

Hypothesis 7. *The concentration-returns relation as well as the effects of higher operating leverage, investment, and operating profitability are ceteris paribus stronger in high aggregate return periods.*

Meanwhile, the secular rise in market concentration and markups due to technological changes that have increased fixed-to-marginal cost ratios leads to the prediction:

Hypothesis 8. *The effect of higher operating leverage on the concentration-returns relation will ceteris paribus become stronger in recent years.*

The increasing dominance of highly efficient firms in recent years implies that (in high entry-threat industries) the importance of large capacities for entry deterrence will be weakened.

Hypothesis 9. *The effect of higher investment on the concentration-returns relation in capital-intensive high entry-threat industries will ceteris paribus become weaker in recent years.*

Entry deterrence and time-varying industry trends also yield predictions for capital investment. Cross-sectionally, capital-intensive high entry-threat industries with higher concentration reflect strong MDE and entry deterrence benefits for investment. But increasing market power of super-efficient firms will attenuate the importance of larger capacities for entry deterrence in recent years.

Hypothesis 10. *In capital-intensive high-entry threat industries, capital investment is ceteris parbus positively related to industry concentration. But the relation of industry concentration and capital investment will become weaker in recent years.*

6 Empirical tests

6.1 Data and sample construction

To test the empirical hypotheses, we obtain data on stock returns from CRSP (Center for Research on Security Prices). We construct a number of firm-specific variables such as size, leverage, etc., using data from Compustat. We employ the text-based TNIC HHI (THHI) industry concentration measure of Hoberg and Phillips (2016), obtained from Gerard Hoberg's website.[28] We merge the quarterly Compustat data with the THHI data.

Our sample begins in 1989 due to the availability of the industry concentration measure. We merge the annual industry concentration measure with the quarterly Compustat files, holding the industry concentration constant over the quarter. To base our empirical tests on the most recent information available to investors, we lag all Compustat data and the industry concentration measure by one quarter. We merge the quarterly Compustat and the industry concentration data with the CRSP monthly data. The Compustat variables are held constant over quarter when we combine the files. We also include the data on Capital Asset Pricing Model (CAPM) betas obtained from the WRDS (Wharton Research and Database Services) betasuite. We drop financial firms from our analysis (SIC codes between 6000 and 7000). Further details on sample construction are provided in the Data appendix. Our final sample covers 1989-2019, with 8784 unique firms. We provide details regarding the computation/construction of empirical measures for the dependent and independent variables (used in our tests) in the Data appendix.

6.2 Results

To test the empirical hypotheses, we run Fama and MacBeth (FM) (1973) cross-sectional regressions using interactions of industry THHI with firm characteristics. We extend the Fama and French (FF) (1992, 1993) three-factor model (FF3) to include financial leverage, similar to Bustamente and Donangelo (2017), since our theoretical model considers unlevered returns. We implement the FF3 model at quarterly frequency, with lagged stock beta (computed every quarter using past 5 years of

[28](http://hobergphillips.tuck.dartmouth.edu/industryconcen.htm). The THHI measure is especially appropriate in our setting because we define industries from the viewpoint of product homogeneity or proximity (see Section 2.1). The THHI measure is derived from text-based descriptions of product characteristics and hence industries are defined in terms of product homogeneity rather than commonality of activities and resource use, which is the basis for conventional industry definitions (see Hoberg and Phillips 2016).

monthly returns).[29] Our basic regression specification is

$$R_{it} = \beta_{0t} + \beta_{1t} THHI_{t-1} + \beta_{2t} X_{i,t-1} + \beta_{3t}(THHI_{t-1} * X_{i,t-1}) + G(controls_t) + \epsilon_{it}, \qquad (23)$$

where X_i is the implicated characteristic, such as OL, OP, and investment. In the standard fashion for FM regressions, the coefficients are estimated each quarter and averaged over time, and the table report the mean coefficients $\bar{\beta} = \text{Avg}(\beta_t)$. The first parts of Hypotheses 1-3 impose restrictions on the sign of β_3, the coefficient of $\frac{\partial}{\partial X}\left(\frac{\partial R}{\partial THHI}\right)$. The second parts of these hypotheses impose restrictions on the sign of the slope or total effect of concentration on returns, $\frac{\partial R}{\partial THHI} = \bar{\beta}_1 + \bar{\beta}_3 \bar{X}$.

Table 6 presents tests of Hypotheses 1-3. The first row shows only the loadings on FF3 and leverage. Consistent with Fama and French (1992), there are strong effects of size and book-to-market (BTM) factors but the market beta is not priced. In the second row, we add industry concentration (THHI), which is not significantly related to equity returns, consistent with the ambiguous evidence in the literature. The third row analyzes the effects of firms' OL on the concentration-returns relation. Similar to the literature (Novy-Marx 2011), we measure OL by the ratio of selling, general, and marketing expense (*xsgaq*) to the sum of *xsagq* and cost of goods sold (*cogsq*). We use OL lagged by a quarter, in accord with the lags of the other covariates. Stock returns are significantly positively related to OL, as is the case in the literature (Lev 1974; Novy-Marx 2011). But the estimated coefficient for THHI × OL ($\bar{\beta}_3$) is negative and statistically significant, indicating that higher OL has a negative impact on the on the concentration-returns relation, supporting the first part of Hypothesis 1. The effect of higher OL is also economically significant. The sample mean of OL is 0.32. Hence, a 1% deviation from the mean OL amplifies the negative effect of concentration on equity returns (that is, $\frac{\partial R}{\partial THHI} = \bar{\beta}_1 + \bar{\beta}_3 O\bar{L}$) by 1.6%. Finally, since $\bar{\beta}_1$ is insignificant but $\bar{\beta}_3$ is reliably negative, the returns-concentration relation is negative for levels of OL above the sample mean.

We test Hypothesis 2 in the next row of Table 6 by analyzing the effects of lagged operating profitability (OP) on the relation of concentration and equity returns. We measure (quarterly) OP as the (total revenue (*revtq*) − *cogsq* − *xsgq*))/(market equity). There is a significant positive relation of OP and returns, consistent with the literature (Fama and French 2015). The estimate of THHI × OP ($\bar{\beta}_3$) is positive and significant, which is consistent with the first part of Hypothesis 3.

[29] While FF (1992) employ portfolio betas for greater precision, using directly computed stock betas should have a minimal impact because (as in FF (1992)) the betas are not significant in any specification. We use WRDS betasuiyte to obtain end-of-quarter betas computed using sixty months returns. We need at least thirty months of returns data to compute betas.

Furthermore, because $\bar{\beta}_3 > 0$ and $\bar{\beta}_1 < 0$, the concentration-returns relation is positive for all firms with (lagged) OP higher than $-(\bar{\beta}_1/\bar{\beta}_3) = 6.4\%$. On average, 12% of firms in a given quarter have lagged OP levels exceeding this threshold (6.4%), which quantifies the second part of Hypothesis 2.

For testing Hypothesis 3, similar to the literature, we measure investment by capital investment-intensity (CII): the ratio of capital investment (*capxq*) and total assets (*atq*). The next row in Table 6 analyzes the effects of firms' one-quarter lagged investment on the concentration-returns relation. Investment loads significantly negatively on average returns, consistent with the literature (Aharoni et al. 2013; Fama and French 2015). For our overall sample, the coefficient of THHI $\times$ CII is positive, consistent with Hypothesis 3, but is not significant. The concentration-returns relation is positive for all firms with (lagged) quarterly CII higher than 3.8%. On average, firms in the top decile of CII (about 7%) in a given quarter have lagged investment exceeding this threshold.

We test Hypothesis 4 using inter-industry variation in $\dot{\delta}_{ind}(u)$, the estimated slope of depreciation with capacity utilization in an industry. We use BEA data on annual depreciation rates and data on annual capacity utilization available from FRB. The industry coverage of capacity utilization data is relatively small compared with the depreciation data. There are twenty four three digit NAICS industries (almost all in manufacturing) where we have both capacity utilization and depreciation data. For each of these industries, we estimate $\dot{\delta}_{ind}(u)$ by regressing depreciation rates on capacity utilization. We classify industries with higher depreciation sensitivity to capacity utilization as the above-median (top-twelve) industries in terms of the estimated slopes $\dot{\delta}_{ind}(u)$, identified by the dummy variable $\mathbb{I}_{\dot{\delta}_{ind}(u)>\text{med}(\dot{\delta}_{ind}(u))}$. We examine the concentration-returns relation in high depreciation sensitivity industries through the interaction term THHI $\times$ $\mathbb{I}_{\dot{\delta}_{ind}(u)>\text{med}(\dot{\delta}_{ind}(u))}$. Hypothesis 4 predicts that the coefficient for this term should be negative. Panel A of Table 7 shows that this coefficient is indeed negative and significant, supporting Hypothesis 4.

To test Hypothesis 5, we identify high capital-intensity and high entry-threat industries using data from Bureau of Labor Studies (BLS) and entry activity analysis in Dunne et al. (1988), respectively. The list of industries is given in the Data appendix. Panel B of Table 7 shows that, compared to the full-sample results in Table 6, the effect of higher CII on the concentration-returns relation—that is, $\frac{\partial}{\partial AG}\left(\frac{\partial R}{\partial THHI}\right)$—increases in statistical significance, consistent with Hypothesis 5.

For testing Hypothesis 6, we identify high entry-threat industries (both capital- and labor-intensive) where large incumbent size deters entry. The US Census Bureau's Survey of Small Business provides data on start-up and business acquisition costs, which is a measure of entry costs (ζ). Retail trade, wholesale trade, and transportation and warehousing have low average startup and, hence, entry costs. However, large incumbent companies in these industries discourage entry (Gershon 2013). Panel C of Table 7 shows that the effects of higher OL are not significant in these industries, in contrast to the full sample results in Table 6, supporting Hypothesis 6.

Now, the high entry-threat industries utilized in Panel C are static over our sample. For robustness, we also construct a time-varying measure of entry threat based on new aggregate business formation quarterly data available from the Bureau of Labor Statistics (BLS) from 1993Q2 onwards. We define high entry quarters as those with rates of new business growth above the sample median growth rate. Panel D of Table 7 shows that the effect of OL on the relation of concentration and returns is not significant prior to high entry quarters, similar to Panel C.

We now examine time-variation in the relation of industry concentration and equity returns. Panels A and B of Table 8 test Hypothesis 7 by analyzing the concentration-returns relation in "boom" periods (Panel A)—defined as periods with quarterly returns exceeding 10%, which holds for about 15% of our sample (19/124 quarters)—and in "no boom" periods (Panel B), that is, the 85% of sample periods with quarterly returns less than 10%. The first row of Panel A shows a significant negative relation of concentration and equity returns during boom periods. In contrast, in Panel B, there is no significant effect of concentration on returns during the no-boom periods. Furthermore, we find that the economic and statistical significance of the effects of OL, CII and OP on the concentration-returns relation is higher during boom quarters relative to "no boom" quarters and therefore the entire sample (Table 6). The results in Panels A and B thus support Hypothesis 7.

We test Hypothesis 8 by considering the effects of high OL on the concentration-returns relation in the second half of our sample, namely, 2005-2019. To do so, we utilize dummy variables $\mathbb{I}_{\geq 2005}$ and $\mathbb{I}_{<2005}$ so that $\bar{\beta} = \bar{\beta}_{t\geq 2005}\mathbb{I}_{\geq 2005} + \bar{\beta}_{t<2005}\mathbb{I}_{<2005}$. Panel C of Table 8 reports coefficients $\bar{\beta}_{t\geq 2005}$. Consistent with the hypothesis, the coefficient of THHI $\times$ OL is significantly negative. For Hypothesis 9, we consider the effects of capital investment on the concentration-returns relation in capital-intensive high entry-threat industries (used in Panel C of Table 7) for the second half of our sample (2005-2019) adopting the dummy variable approach outlined above. Panel C of Table 8 also shows that the coefficient of THHI $\times$ CII in the returns regression for the annunciated industries is insignificant during 2005-2019, supporting Hypothesis 9.

For testing Hypothesis 10, we utilize a panel regression of CII on lagged THHI, identifying capital intensive high entry-threat industries through a dummy variable (High Entry), and controlling for (lagged) Tobin's Q, size, and leverage. We follow standard practice in the literature and use the book values of firms' assets as proxies for their replacement value, and compute Q as the ratio of the sum of market value of equity and total liabilities and the sum of book value of equity and total liabilities. Table 9 shows that the coefficient for THHI $\times$ High Entry is positive and significant, supporting the first part of the Hypothesis. We test for weakening of this effect in recent years through the dummy variable (THHI $\times$ High Entry)$\mathbb{I}_{\geq 2005}$. Indeed, the coefficient for this term is negative and significant, consistent with the second part of the Hypothesis.

7 Summary and Conclusions

Oligopolies are ubiquitous and real and financial outcomes of strategic interaction by firms in such industries attract long-standing, substantial attention. In a dynamic general equilibrium production economy setting with oligopolistic sectors and capital investment, our analysis theoretically, quantitatively and empirically highlights the interaction of firm- and aggregate-level real outcomes with financial markets. In response to productivity shocks, industry firms strategically choose capacity utilization and investment to maximize tacit-collusive equity values subject to a no-defection constraint governed by on- and off-equilibrium equity values, thereby generating substantial interaction between real outcomes and financial markets. Quantitatively, with moderate risk-aversion and standard calibration, the model generates relatively high equity premium, low real risk-free rate, as well as industry and aggregate fluctuations and autocorrelations that help explain the data. Strategic interaction in our setting also generates novel perspectives on the role of capital and capital-related frictions—such as, fixed and capital adjustment costs, endogenous depreciation, and entry capacity costs—on priced risks. We find empirical support for novel predictions regarding the effects of firm (operating leverage, operating profits and investment) and industry (durability and entry threat) characteristics, as well time-varying aggregate financial market conditions, on the cross-sectional relation of industry concentration and returns.

Our analysis highlights the significant effects of capital and real frictions on equilibrium outcomes with strategic interaction in dynamic oligopolies. For tractability, we do not consider financial leverage; therefore, the effects of capital and operating leverage we find could reflect the latent impact of leverage. Analyzing tacit collusion in a dynamic production-based asset pricing model with endogenous operational and financial leverage is an important area for future research.

Data appendix

D1. Variable definitions

Book Equity (BE): We define $book\ equity = shareholder\ equity + txditcq - pstkq$, where $txditcq$ are the (quarterly) deferred taxes and investment tax credit and $pstkq$ is the total preferred stock. We use Compustat variable $seqq$ for shareholder equity. If it is not available, we compute shareholder equity as the sum of total common equity ($ceqq$) and $pstkq$. If this is not available, we compute shareholder equity as the difference between total assets (atq) and total liabilities (ltq). We set negative BE to "not available" in our tests.

Boom periods: We define boom as periods with quarterly market returns exceeding 10%, which holds for about 15% of our sample—and "no boom" periods (Panel B), that is, the 85% of sample periods with monthly market returns less than 10%.

Capital investment intensity (CII): Defined as the ratio of capital investment (capxq) and the total assets (atq).

Industry concentration: As our industry concentration measure, we utilize the Text-based Network Industry Classifications HHI (THHI) measure of Hoberg and Phillips (2016).

Industry entry threat: In Panel C of Table 7 and in Table 9, we identify capital-intensive high entry threat industries through a dummy variable (at the two-digit SIC code level) based on industry-specific entry activity documented in Dunne et al. (1988) and capital intensity in Kutscher and Mark (1983). These industries include printing, chemicals, business services, research, and engineering and management services. In Panel D of Table 7, the high entry-threat industries (at the two digit SIC code level) are retail trade, wholesale trade, and transportation and warehousing.

Leverage: Computed as the ratio of total liabilities (ltq) and the sum of total liabilities and market equity.

Market equity (ME): Computed as the number of shares outstanding times the stock price.

Operating leverage (OL): Computed as the ratio of Selling, General, and Marketing expense (xsgaq) and the sum of Selling, General, and Marketing expense and Cost of Goods sold (cogsq).

Operating profitability (OP): Computed as the ratio: (Total Revenue (REVTQ) - Cost of Goods Sold (cogsq) - Selling, General, and Marketing expense (xsga))/Market Equity (ME).

Stock returns: These are firm-level quarterly total equity returns, measured in percentages.

Tobin's Q: We follow standard practice in the literature and use the book values of firms' assets as proxies for their replacement value, and compute Q by dividing the sum of market value of equity and total liabilities with the sum of book value of equity and total liabilities.

D2. Sample construction

For dependent variables that use Compustat quarterly data, we drop firms with irregular fiscal quarter-ends, that is, with quarter-ends other than March, June, September, and year-end. To reduce the likelihood of returns being driven by outliers, we winsorize the compustat data (e.g., markups and capex) at 1 percentile and 99 percentile levels of their distributions. We then identify firms for which there is a THHI measure. The sample selection procedure results in 8784 unique firms.

References

Abel, A. 1981. A dynamic model of investment and capacity utilization. *Quarterly Journal of Economics* 96:379-404.

Abreu, D. 1986. Extremal equilibria of oligopolistic supergames. *Journal of Economic Theory* 39:191-225.

Aguerrevere, F. 2009. Real options, product market competition, and asset returns. *Journal of Finance* 64:957–83.

Aharoni, G., B. Grundy, and Q. Zeng. 2013. Stock returns and the Miller Modigliani valuation formula: Revisiting the Fama French analysis. *Journal of Financial Economics* 110:347-357.

Andreasen, M., J. Fernandez-Villaverde, and J. Rubio-Ramirez. 2017. The pruned state-space system for non-linear DSGE models: Theory and empirical applications. *Review of Economic Studies* 81:1-49.

Autor, D., D. Dorn, L. Katz, C. Patterson, and J. Van Reenen. 2020. The fall of the labor share and the rise of superstar firms. *Quarterly Journal of Economics* 135:645–709.

Bansal, R., and A. Yaron. 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59:1481-1509.

Baxter, A., and Curcini. 1993. Explaining saving-investment correlations. *American Economic Review* 83:416-436.

Benoit, J.P., and V. Krishna. 1987. Dynamic duopoly: Prices and quantities. *Review of Economic Studies* 54:21–35.

Berk, J., R. Green, and V. Naik. 2004. Valuation and return dynamics of new ventures. *Review of Financial Studies* 17:1-35.

Berry, S., M. Gaynor, and F. Morton. 2019. Do increasing markups matter? Lessons from empirical industrial organization. *Journal of Economic Perspectives* 33:44-68.

Bilbiie, F., F. Ghironi, and M. Melitz. 2012. Entry, product variety, and business cycles. *Journal of Political Economy* 120:304-345.

Blume, L., D. Easley, and M. O'Hara. 1982. Characterization of optimal plans for stochastic dynamic programs. *Journal of Economic Theory* 28:221-234.

Boldrin, M., L. Christiano, and M. Eichenbaum. 2001. Habit persistence, asset returns, and the business cycle. *American Economic Review* 91:149-166.

Bond, P., A. Edmans, and I. Goldstein. 2012. The real effects of financial markets. *Annual Review of Financial Economics* 4:339-360.

Brock, W., 1972. On models of expectations that arise from maximizing behavior of economic agents over time. *Journal of Economic Theory* 5:348-376.

Brock, W., 1982. Asset pricing in production economies, In *The Economics of Information and Uncertainty* ed. D. J. McCall, 1–46. University of Chicago Press: Chicago.

Brock, W., and J. Scheinkman. 1985. Price setting supergames with capacity constraints. *Review of Economic Studies* 52:371–382.

Broda, C., and D. Weinstein. 2006. Globalization and the gains from variety. *Quarterly Journal of Economics* 121:541-86.

Bustamente, C. 2015. Strategic investment and industry risk dynamics. *Review of Financial Studies* 28:297-314.

Bustamante, C., and A. Donangelo. 2017. Product market competition and industry returns. *Review of Financial Studies* 30:4216-4266.

Caldera, D., J. Fernandez-Villaverde, J. Rubio-Ramirez, and W. Yao. 2012. Computing DSGE models with recursive preferences and stochastic volatility. *Review of Economic Dynamics* 15:188-206.

Campbell, J., and R. Shiller. 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1:195-228.

Carlson, M., A. Fisher, and R. Giammarino. 2004. Capital investment and asset price dynamics: Implications for the cross-section of returns. *Journal of Finance* 59:2577-2603.

Chari, V., P. Kehoe, and E. McGratten. 2000. Sticky price models of the business cycle: Can the contract multiplier solve the persistence problem? *Econometrica* 68:1151-1179.

Chen, Q., I. Goldstein, and W. Jiang. 2007. Price informativeness and investment sensitivity to stock price. *Review of Financial Studies* 20:619-650.

Cochrane, J. 1991. Production-based asset pricing and the link between stock returns and economic fluctuations. *Journal of Finance* 46:209–237.

Cooper, R., and J. Haltiwanger. 2006. On the nature of capital adjustment costs. *Review of Economic Studies* 73:611-633.

Corhay, A., H. Kung, and L. Schmid. 2020. Competition, markups, and predictable returns. *Review of Financial Studies* 33:5906-5939.

Cournot, A. 1838. *Recherches sur les principles mathematiques de la theorie des richesses.* Paris.

Croce, M. 2014. Long-run productivity risk: A new hope for production-based asset pricing? *Journal of Monetary Economics* 66:13-31.

De Loecker, J., J. Eckhout, and G. Unger. 2020. The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics* 135:561–644.

Dixit, A. 1980. The role of investment in entry-deterrence. *Economic Journal* 90:95-106.

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

Dou, W., Y. Ji, and W. Wu. 2019. Competition, profitability, and risk premia. *Journal of Financial Economics* 140:582-620.

Dou, W., Y. Ji, and W. Wu. 2021. The oligopoly Lucas tree. *Review of Financial Studies* 35:3867-3921.

Duffee, G. 2005. Time variation in the covariance between stock returns and consumption growth. Journal of Finance 60:1673-1712.

Dunne, T., M. Roberts, and L. Samuelson. 1986. Patterns of firm entry and exit in U.S. manufacturing industries. *Rand Journal of Economics* 19: 495-515.

Epstein, L., and S. Zin. 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica* 57:937-969.

Fama, E., and K. French. 1992. The cross-section of expected stock return. *Journal of Finance* 47:427-456.

Fama, E., and K. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33:3-56.

Fama, E., and K. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116:1-22.

Fama, E., and J. MacBeth. 1973. Risk, return, and equilibrium: empirical tests. *Journal of Political Economy* 81:607-636.

Farhi, E., and F. Gourio. 2018. Accounting for macro-finance trends: Market power, intangibles, and risk premia. *Brookings Papers on Economic Activity* Fall:147-250.

Federal Reserve Board. 2022. Statistical release G.17 (April 15).

Fernandez-Villaverde, J., J. Rubio-Ramirez, and F. Schorfheide. 2016. Solution and estimation methods for DSGE models. In J. Taylor and H. Uhlig (eds.) *Handbook of Macroeconomics,* 530-713.

Foerster, A., P-D Sarte, and M. Watson. 2011. Sectoral versus aggregate shocks: A structural factor analysis of industrial production. *Journal of Political Economy* 119:1-38.

Friedman, J. 1983. *Oligopoly theory.* Cambridge, UK: Cambridge University Press.

Fudenberg, D., and J. Tirole. 1991. *Game theory.* Cambridge, MA: MIT Press.

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

Gârleanu, N., S. Pangeas, and J. Yu. 2012. Technological growth and asset pricing. *Journal of Finance* 67:1265-1292.

Gershon, L. 2013. Study: Industries with low barriers to entry. https://www.creditdonkey.com/barriers-to-entry.html (accessed September 14, 2022).

Gomme, P., and D. Lkhagvasuren, 2013. Calibration and simulation of DSGE models, In Hashimzade, N., and M.. Thornton (eds) *Handbook of Research Methods and Applications in Empirical Macroeconomics*, 575-592, Edward Elgar Publishing.

Greeenwood, J., Z. Hercovitz, and G. Huffman. 1988. Investment, capacity utilization, and the real business cycle. *American Economic Review* 78:402-418.

Gu, L. 2016. Product market competition, R&D investment and stock returns. *Journal of Financial Economics* 119:441-455.

Hansen, L., and K. Singleton. 1982. Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica* 50:1269-1286.

Hansen, L., and K. Singleton. 1983. Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of Political Economy* 91:249-265.

Hoberg, G., G. Phillips. 2016. Text-based network industries and endogenous product differentiation. *Journal of Political Economy* 124:1423-1465.

Horvath, M. 2000. Sectoral shocks and aggregate fluctuations. *Journal of Monetary Economics* 45:69-106.

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

Jaimovich, N., and S. Rebelo. 2009. Can news about the future drive the business cycle? *American Economic Review* 99:1097-1118.

Jermann, U. 1998. Asset pricing in production economies. *Journal of Monetary Economics* 41:257-275.

Jin, H., and K. Judd. 2002. Perturbation methods for general dynamic stochastic models. Discussion paper. Stanford University.

Khan, A., and J. Thomas. 2008. Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics. *Econometrica* 76:395-436.

Kim, J., S. Kim, E. Schaumburg, and C. Sims. 2008. Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models. *Journal of Economic Dynamics and Control* 32:3397-3414.

Kreps, D., and J. Scheinkman. 1983. Quantity precommitment and Bertrand competition yields Cournot outcomes. *Rand Journal of Economics* 14:326–377.

Kreps, D., and R. Wilson. 1982. Sequential equilibria. *Econometrica* 50:863-894.

Kutscher, R., and J. Mark. 1983. The service-producing sector: some common perceptions reviewed. Bureau of Labor Statistics, *Monthly Labor Review*, April:21-24.

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

Lev, B. 1974. On the association between operating leverage and risk. *Journal of Financial and Quantitative Analysis* 9:627-641.

Lewellen, B. 1999. The time-series relations among expected return, risk, and book-to-market. *Journal of Financial Economics* 54:5-43.

Liu, L., T. Whited, and L. Zhang. 2009. Investment-based expected stock returns. *Journal of Political Economy* 117:1105-1139.

Loualiche, E. 2021. Asset pricing with entry and imperfect information. *Journal of Finance,* Forthcoming.

Lucas, R. 1967. Adjustment costs and the theory of supply. *Journal of Political Economy* 75:321-334.

Mehra, R., and E. Prescott. 1985. The equity premium: A puzzle. *Journal of Monetary Economics* 15:145-161.

Morck, R., A. Shleifer, and R. Vishny. 1990. The stock market and investment: Is the market a sideshow? *Brookings Papers on Economic Activity* 21:157-216.

Morellec, E., and A. Zhdanov. 2019. Product market competition and option prices. *Review of Financial Studies* 32:4343-4386.

Novy-Marx, R. 2011. Operating leverage. *Review of Finance* 15:103-134.

Opp, M., C. Parlour, and J. Walden, 2014. Markup cycles, dynamic misallocation, and amplification. *Journal of Economic Theory* 154:126-161.

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

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

Schmitt-Grohé, S., and Uribe, M. 2004. Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of Economic Dynamics and Control* 28:755–775.

Selten, R. 1975. Re-examination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory* 4:25-55.

Siegfried, J., and L. Evans. 1994. Empirical studies of entry and exit: A survey of the evidence. *Review of Industrial Organization* 9: 121-155.

Smith, A. 1776. *An inquiry into the nature and causes of the wealth of nations.* London: W. Strahan.

Spence, M. 1977. Entry, investment, and oligopolistic pricing. *Bell Journal of Economics* 8:534-544.

Veracierto, M. 2002. Plant-level irreversible investment and equilibrium business cycles. *American Economic Review* 92:181-197.

Weil, P. 1989. The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics* 24:401-421.

Zhang, L. 2005. The value premium. *Journal of Finance* 60:67-103.

Table 1. Baseline calibration

This table displays the calibration for simulations of the baseline model.

Consumption and Production	
Annual discount factor (α)	0.99
Risk aversion (γ)	5
Intertemporal elasticity of substitution (η^{-1})	1.9
Intratemporal (product) elasticity of substitution (σ)	5.99
Output elasticity of capita: Concentrated sector (ψ_1)	0.49
Output elasticity of capital: Low-Concentration sector (ψ_2)	0.40
Marginal cost: Concentrated sector (h_1)	0.1
Marginal cost: Low-Concentration sector (h_2)	0.1
Fixed cost: Concentrated sector (o_1)	0.01
Fixed cost: Low-Concentration sector (o_2)	0.45
Capital adjustment cos: Concentrated sector (φ_1)	8
Capital adjustment cost: Low-Concentration sector (φ_2)	1
Annual depreciation rate: Concentrated sector (δ_1)	5%
Annual depreciation rate: Low-Concentration sector (δ_2)	5.3%
Variance-Covariance Matrix of Productivity Shocks (Annualized)	
Volatility of Concentrated sector shock ε_1 (λ_1)	0.1%
Volatility of Low-concentration sector shock ε_2 (λ_2)	1.5%
Correlation of ε_1 and ε_2 (λ_{12})	-0.07
Autocorrelation Coefficients (Annual)	
Coefficient for ε_1 (ρ_1)	0.73
Coefficient for ε_2 (ρ_2)	0.96

Table 2. Main real moments (baseline model)

This table presents the moments of salient firm-level and aggregate real variables from simulations of the baseline model with eight firms in the concentrated sector and thirty eight firms in the low-concentration sector. The displayed moments are based on 100,000 simulations each with 59 yearly observations.

Variable	Model	Data
Panel A: Standard Deviations (%)		
Capacity Utilization (u_1)	9.35	18.11
Log Changes		
Investment Rate: Concentrated sector (g_{ir_1})	8.88	15.69
Investment Rate: Low-Concentrated sector (g_{ir_2})	0.73	14.27
Aggregate Real Consumption (g_c)	2.67	2.56
Aggregate Financial Income (W)	3.4	6.29
Aggregate Price Index (P)	0.73	2.8
Panel B: Correlations		
Corr(IR_1, W)	0.05	0.65
Corr(IR_1, P)	0.03	0.8
Cov(u_1, W)	-0.92	-0.57
Cov(u_1, P)	-0.95	-0.58
Panel C: Autocorrelation coefficients (Log changes)		
Investment Rate: Concentrated sector (g_{ir1})		
ACF(1)	-0.2	-0.1
ACF(2)	-0.12	-0.04
ACF(3)	-0.07	0.01
ACF(4)	-0.04	-0.07
Aggregate Consumption (g_c)		
ACF(1)	-0.01	0.25
ACF(2)	-0.01	0.05
ACF(3)	-0.01	-0.02
ACF(4)	-0.003	-0.09
Aggregate Income (g_w)		
ACF(1)	-0.01	0.52
ACF(2)	-0.01	0.18
ACF(3)	-0.01	0.12
ACF(4)	-0.001	0.16

Table 3. Main financial moments (baseline model)

This table presents the moments of salient firm-level and aggregate financial variables from simulations of the baseline model with eight firms in the concentrated sector and thirty eight firms in the low-concentration sector. The displayed moments are based on 590,000 yearly simulations (100,000 simulations of our sample of 59 years).

Panel A: Mean and Standard Deviations (%)

	Mean		S.D. (Log Changes)	
	Model	**Data**	**Model**	**Data**
Real Risk-Free Rate (r_f^r)	0.86	0.91	0.11	2.15
Risk-Premium: Concentrated sector (ERP_1)	2.87	6.94	11.37	22.95

Panel B: Autocorrelations

	Model	**Data**
Real Risk-Free Rate (r_f^r)		
ACF(1)	0.96	0.73
ACF(2)	0.93	0.55
ACF(3)	0.92	0.42
ACF(4)	0.91	0.21
Risk-Premium: Concentrated sector (ERP_1)		
ACF(1)	0.88	-0.06
ACF(2)	0.8	-0.25
ACF(3)	0.75	0.01
ACF(4)	0.71	0.03

Table 4. Main moments with entry and exit

This table presents the moments of salient firm-level and aggregate variables from simulations of the model with entry and exit. The displayed moments are based on 590,000 yearly simulations (100,000 simulations of our sample of 59 years).

Panel A: Standard Deviations (%)				
Variable	**Model**	**Data**		
Capacity Utilization (u_1)	28.09	18.11		
Log Changes				
Investment Rate: Concentrated sector (g_{ir1})	19.68	15.69		
Aggregate Consumption (g_c)	5.72	2.56		
Aggregate Financial Income (W)	3.76	6.29		
Aggregate Price Index (P)	2.80	2.89		
Panel B: Autocorrelation Coefficients (Log Changes)				
Aggregate Consumption (g_c)				
ACF(1)	0.18	0.25		
ACF(2)	0.05	0.05		
ACF(3)	0.06	-0.02		
ACF(4)	0.06	-0.09		
Aggregate Income (g_w)				
ACF(1)	0.32	0.52		
ACF(2)	0.11	0.18		
ACF(3)	0.12	0.12		
ACF(4)	0.12	0.16		
Panel C: Mean and Standard of Financial Returns (%)				
	Mean		**S.D. (Log Changes)**	
	Model	**Data**	**Model**	**Data**
Real Risk-Free Rate (r_f^r)	0.94	0.91	2.82	2.15
Risk-Premium: Concentrated sector (ERP_1)	5.19	6.94	32.22	22.95

Table 5. Effects of adjustment costs and sectoral productivity shock processes

This table analyzes the effects of varying adjustment costs and parameters governing the sectoral productivity shock processes (relative to the baseline calibration given in Table 1) on the model with entry and exit (see Table 4). The displayed moments are based on 590,000 yearly simulations (100,000 simulations of our sample of 59 years).

	High φ_1	Low ρ_1	Low λ_1	Low ρ_2	Low λ_2
Panel A: Standard Deviations (%)					
Capacity Utilization (u_1)	26.48	28.09	28.14	23.18	14.85
Log Changes					
Investment Rate: Concentrated sector (g_{ir1})	12.94	19.68	19.63	17.48	12.98
Aggregate Consumption (g_c)	6.03	5.72	5.7	4.92	3.42
Aggregate Financial Income (W)	3.88	3.76	3.75	3.2	2.03
Aggregate Nominal Consumption (PC)	3.88	3.76	3.75	3.2	2.03
Aggregate Price Index (P)	2.98	2.8	2.79	2.4	1.73
Panel B: Mean and Standard of Financial Returns (%)					
Mean					
Real Risk-Free Rate (r_f^r)	0.98	0.94	0.94	0.90	0.98
Risk-Premium: Concentrated sector (ERP_1)	5.99	5.18	5.24	1.00	2.29
Standard Deviation (Log Changes)					
Real Risk-Free Rate (r_f^r)	2.99	2.82	2.8	2.43	1.74
Risk-Premium: Concentrated sector (ERP_1)	31.5	32.22	32.28	26.83	17.38

Table 6. Industry concentration, stock returns and firm characteristics

This table analyzes the effects of operating leverage (OL), operating profitability (OP) and capital investment-intensity (CII) on the relation of industry concentration and stock returns at the firm-level using cross-sectional Fama and MacBeth (1973) regressions. The Text-based Network Industry Classifications Herfindahl-Hirschman Index (THHI) of Hoberg and Phillips (2016) is used as the measure of industry concentration. The definitions of other variables are provided in the Data appendix. The sample period is 1989-2019. The annual industry concentration measure is merged with the quarterly Compustat files, holding the industry concentration constant over the quarter. All Compustat data and the industry concentration measure are lagged by one quarter. Quarterly Compustat and industry concentration data are then merged with CRSP monthly data. The Compustat variables are held constant over quarter when we combine the files. We also include the data on Capital Asset Pricing Model (CAPM) betas obtained from the WRDS (Wharton Research and Database Services) betasuite. Financial firms are dropped (SIC codes between 6000 and 7000). Further details on sample construction are provided in the Data appendix. The sample has 8784 unique firms. For each estimated coefficient, the number in the parentheses report the corresponding Newey-west adjusted standard error. *** indicates significance at 1% level, ** indicates significance at 5% level, and * indicates significance at 10% level.

Beta	ln(ME)	BE/ME	Leverage	THHI	OL	THHI x OL	OP	THHI x OP	CII	THHI x CII
0.61	-0.46***	1.64***	0.14							
(0.50)	(0.16)	(0.47)	(1.42)							
0.55	-0.49***	1.62***	0.12	-0.87						
(0.48)	(0.17)	(0.47)	(1.42)	(0.57)						
0.44	-0.41***	1.70***	1.17	0.56	3.45*	-4.64*				
(0.45)	(0.15)	(0.58)	(1.23)	(0.60)	(1.49)	(2.23)				
0.75	-0.66***	1.42**	-5.65***	-2.34***			62.64***	36.74***		
(0.47)	(0.16)	(0.58)	(1.28)	(0.65)			(5.95)	(11.51)		
0.53	-0.46***	1.75***	0.35	-1.50**					-34.03**	39.69
(0.46)	(0.18)	(0.50)	(1.46)	(0.71)					(13.75)	(24.21)

Table 7. Industry concentration, stock returns and industry characteristics

This table analyzes the relation of industry concentration and stock returns at the firm-level using cross-sectional Fama and MacBeth (1973) regressions for industries with high depreciation costs of capacity utilization (identified with a dummy variable), as well as capital- and labor-intensive high entry-threat industries (see Data appendix and Section 6.3). The Text-based Network Industry Classifications Herfindahl-Hirschman Index (THHI) of Hoberg and Phillips (2016) is used as the measure of industry concentration. The definitions of other variables are provided in the Data appendix. The sample period is 1989-2019. The annual industry concentration measure is merged with the quarterly Compustat files, holding the industry concentration constant over the quarter. All Compustat data and the industry concentration measure are lagged by one quarter. Quarterly Compustat and industry concentration data are then merged with CRSP monthly data. The Compustat variables are held constant over quarter when we combine the files. We also include the data on Capital Asset Pricing Model (CAPM) betas obtained from the WRDS (Wharton Research and Database Services) betasuite. Financial firms are dropped (SIC codes between 6000 and 7000). Further details on sample construction are provided in the Data appendix. The sample has 8784 unique firms. For each estimated coefficient, the number in the parentheses report the corresponding Newey-west adjusted standard error. *** indicates significance at 1% level, ** indicates significance at 5% level, and * indicates significance at 10% level.

Beta	ln(ME)	BE/ME	Leverage	THHI	THHI x $\mathbb{I}_{\dot{\delta}_{ind}(u)>\text{med}(\dot{\delta}_{ind}(u))}$	OL	THHI x OL	CII	THHI x CII
Panel A: High Depreciation Slope									
0.62	-0.43**	2.83***	0.31	-2.40***	-10.67*				
(0.50)	(0.18)	(0.78)	(1.70)	(0.93)	(6.04)				
Panel B: High Entry Threat (Capital-Intensive Industries)									
0.04	-0.61***	0.73	2.90***	-2.33***				-58.99*	74.85*
(0.42)	(0.26)	(0.54)	(1.84)	(1.27)				(20.81)	(44.28)
Panel C: High Entry Threat (General)									
-2.37	-0.46	-0.81	-0.44	19.71		18.63	-75.16		
(2.53)	(0.33)	(1.33)	(0.53)	(19.2)		(18.52)	(72.9)		
Panel D: High Entry Threat (Time-varying, General)									
0.19	-0.34	2.19*	0.53	0.80		3.01*	-3.66		
(0.61)	(0.23)	(1.26)	(1.81)	(0.90)		(1.57)	(2.39)		

Table 8. Time-variation in effects of industry concentration on stock returns

This table analyzes time-variation in the relation of industry concentration and stock returns at the firm-level. The Text-based Network Industry Classifications Herfindahl-Hirschman Index (THHI) of Hoberg and Phillips (2016) is used as the measure of industry concentration. The definitions of other variables are provided in the Data appendix. The sample period is 1989-2019. The annual industry concentration measure is merged with the quarterly Compustat files, holding the industry concentration constant over the quarter. All Compustat data and the industry concentration measure are lagged by one quarter. Quarterly Compustat and industry concentration data are then merged with CRSP monthly data. The Compustat variables are held constant over quarter when we combine the files. We also include the data on Capital Asset Pricing Model (CAPM) betas obtained from the WRDS (Wharton Research and Database Services) betasuite. Financial firms are dropped (SIC codes between 6000 and 7000). Further details on sample construction are provided in the Data appendix. The sample has 8784 unique firms. For each estimated coefficient, the number in the parentheses report the corresponding Newey-west adjusted standard error. *** indicates significance at 1% level, ** indicates significance at 5% level, and * indicates significance at 10% level.

Beta	ln(ME)	BE/ME	Leverage	THHI	OL	THHI x OL	OP	THHI x OP	CII	THHI x CII
Panel A: Boom Months										
7.18***	-1.44***	1.62	0.75	-4.70**						
(1.47)	(0.54)	(1.06)	(4.81)	(1.96)						
6.35***	-1.05**	1.56	6.97*	3.11	16.50**	-24.33**				
(1.30)	(0.45)	(1.03)	(3.89)	(2.51)	(6.62)	(10.95)				
6.95***	-1.41***	1.24	-2.18	-8.21***			31.31*	89.55***		
(1.42)	(0.50)	(1.14)	(4.12)	(2.22)			(15.60)	(30.37)		
6.87***	-1.30**	2.03*	1.15	-6.83***					-112.48***	137.99**
(1.42)	(0.53)	(1.08)	(4.86)	(2.62)					(36.16)	(69.24)
Panel B: No-Boom Months										
-0.64	-0.31**	1.62***	-0.00	-0.18						
(0.40)	(0.15)	(0.51)	(1.23)	(0.45)						
-0.63	-0.30**	1.72***	0.12	0.10	1.09	-1.08				
(0.40)	(0.14)	(0.65)	(1.15)	(0.58)	(1.04)	(1.51)				
-0.37	-0.52***	1.45**	-6.28***	-1.28**			68.30***	27.19**		
(0.41)	(0.14)	(0.64)	(1.19)	(0.53)			(6.18)	(12.26)		
-0.62	-0.31**	1.70***	0.20	-0.54					-19.84	21.90
(0.39)	(0.15)	(0.54)	(1.26)	(0.52)					(13.44)	(23.10)
Panel C: Data from 2005 Onwards										
0.25	-0.14	0.66	2.13	1.40**	2.05**	-4.40***				
(0.53)	(0.11)	(0.58)	(1.54)	(0.64)	(0.96)	(1.62)				
-0.17	-0.17	0.40	3.77**	-1.11					-50.84**	26.47
(0.52)	(0.22)	(0.74)	(1.79)	(1.10)					(27.63)	(61.73)

Table 9. Entry deterrence and capital investment

This table analyzes effects of strategic entry deterrence on capital investment-intensity (CII) and the time-variation in these effects. The Text-based Network Industry Classifications Herfindahl-Hirschman Index (THHI) of Hoberg and Phillips (2016) is used as the measure of industry concentration. The definitions of other variables are provided in the Data appendix. The sample period is 1989-2019. The annual industry concentration measure is merged with the quarterly Compustat files, holding the industry concentration constant over the quarter. All Compustat data and the industry concentration measure are lagged by one quarter. Quarterly Compustat and industry concentration data are then merged with CRSP monthly data. The Compustat variables are held constant over quarter when we combine the files. We also include the data on Capital Asset Pricing Model (CAPM) betas obtained from the WRDS (Wharton Research and Database Services) betasuite. Financial firms are dropped (SIC codes between 6000 and 7000). Further details on sample construction are provided in the Data appendix. The sample has 8784 unique firms. For each estimated coefficient, the number in the parentheses report the corresponding Newey-west adjusted standard error. For ease of tabulation, the dependent variable is scaled by 100. *** indicates significance at 1% level, ** indicates significance at 5% level, and * indicates significance at 10% level.

Q	ln(ME)	Leverage	THHI	High Entry	THHI x High Entry	(THHI x High Entry)$\mathbb{I}_{\geq 2005}$
0.00	0.03***	-1.06***	-1.09***	-1.18***	1.21**	
(0.00)	(0.01)	(0.06)	(0.06)	(0.04)	(0.07)	
Q	ln(ME)	Leverage	THHI	High Entry	THHI x High Entry	(THHI x High Entry)$\mathbb{I}_{\geq 2005}$
0.00	0.03***	-1.06***	-1.09***	-1.18***	1.26**	-0.12*
(0.00)	(0.01)	(0.06)	(0.06)	(0.04)	(0.07)	(0.07)

Figure 1. Responses to concentrated sector productivity shock (baseline model)

This figure displays ten-year impulse response functions of salient firm-level and aggregate variables in the baseline model to a one standard deviation shock of capital productivity in the concentrated sector.

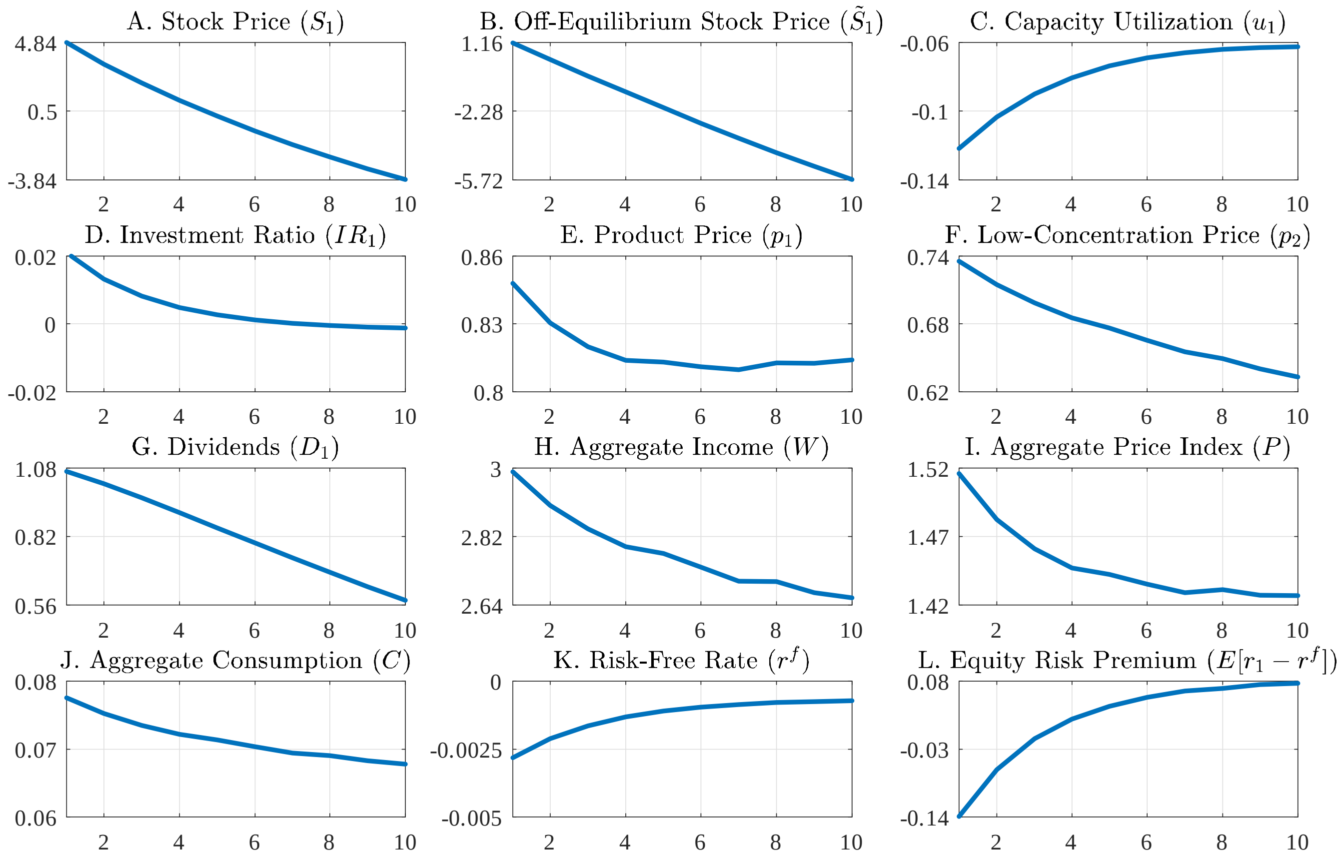


Figure 2. Responses to low-concentration sector productivity shock (baseline model)

This figure displays ten-year impulse response functions of salient firm-level and aggregate variables in the baseline model to a one standard deviation shock of capital productivity in the low-concentration sector.

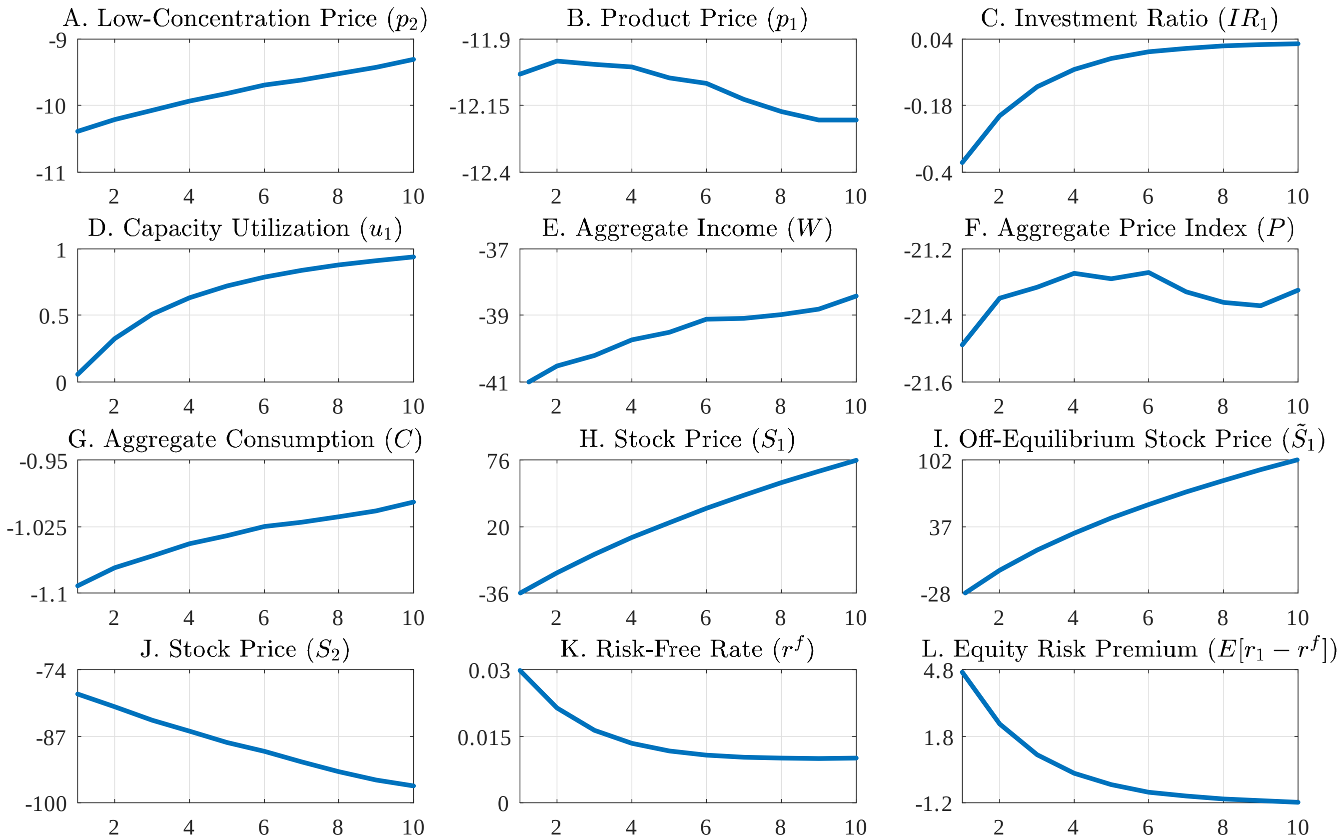


Figure 3. Responses to concentrated sector productivity shock (with entry and exit)

This figure displays ten-year impulse response functions of salient firm-level and aggregate variables in the model with entry and exit to a one standard deviation shock of capital productivity in the concentrated sector .

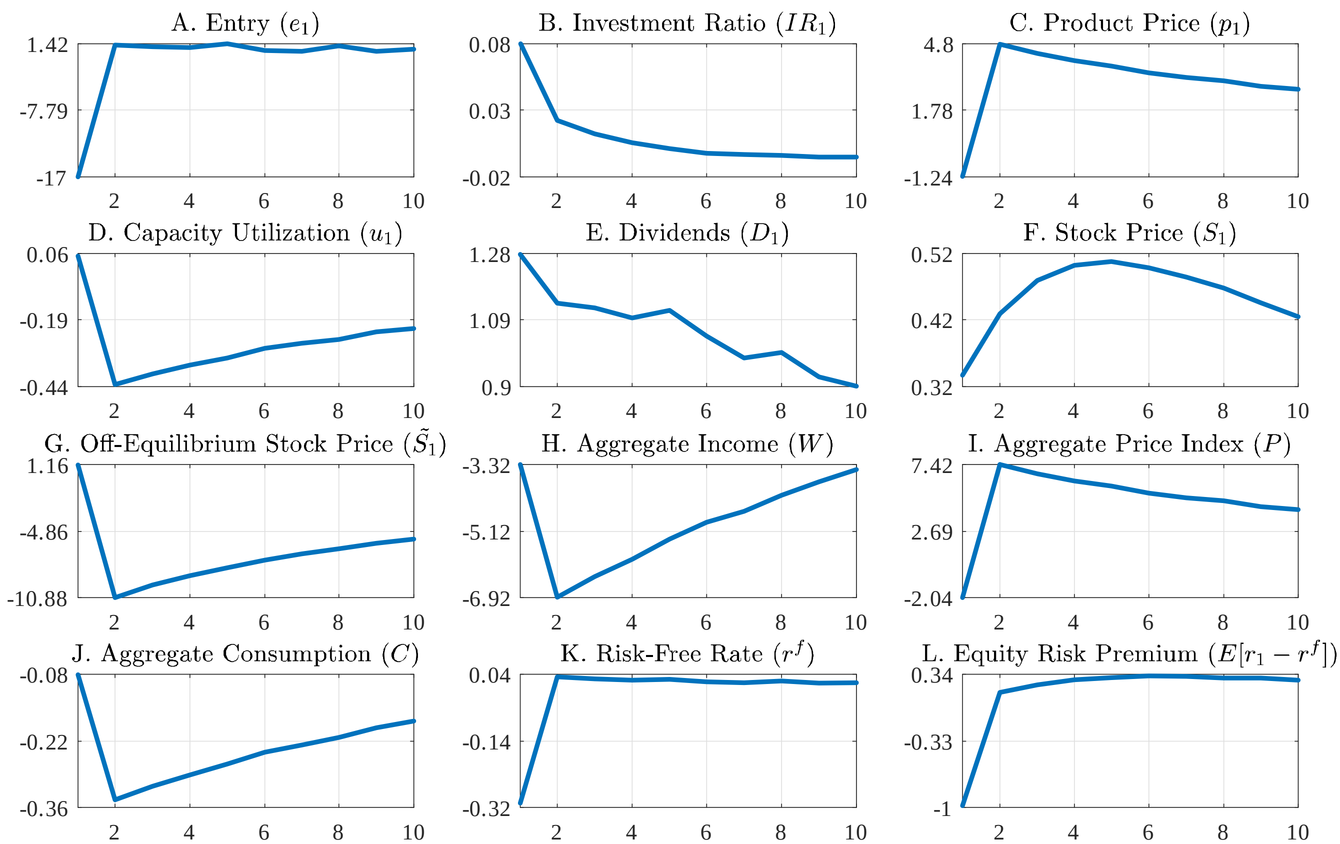


Figure 4. Industry concentration and equity returns with high operating leverage

This figure displays, through a three-dimensional surface plot, the relation of mean equity returns and industry concentration (inversely related to the number of industry firms) for high ratios of fixed-to-marginal costs, or high (operating leverage/operating profits) ratios, in the industry equilibrium analyzed in Section 5.1. The mean returns are computed from 1000 simulations of 59 yearly observations.

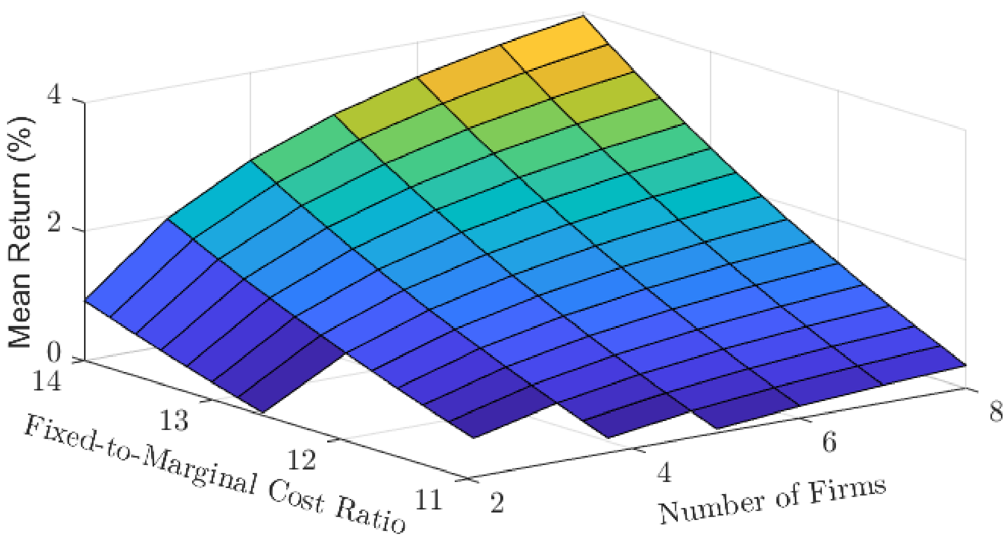


Figure 5. Industry concentration and equity returns with high operating profits

This figure displays, through a three-dimensional surface plot, the relation of mean equity returns and industry concentration (inversely related to the number of industry firms) for low ratios of fixed-to-marginal costs, or high (operating profits/operating profits) ratios, in the industry equilibrium analyzed in Section 5.1. The mean returns are computed from 1000 simulations of 59 yearly observations.

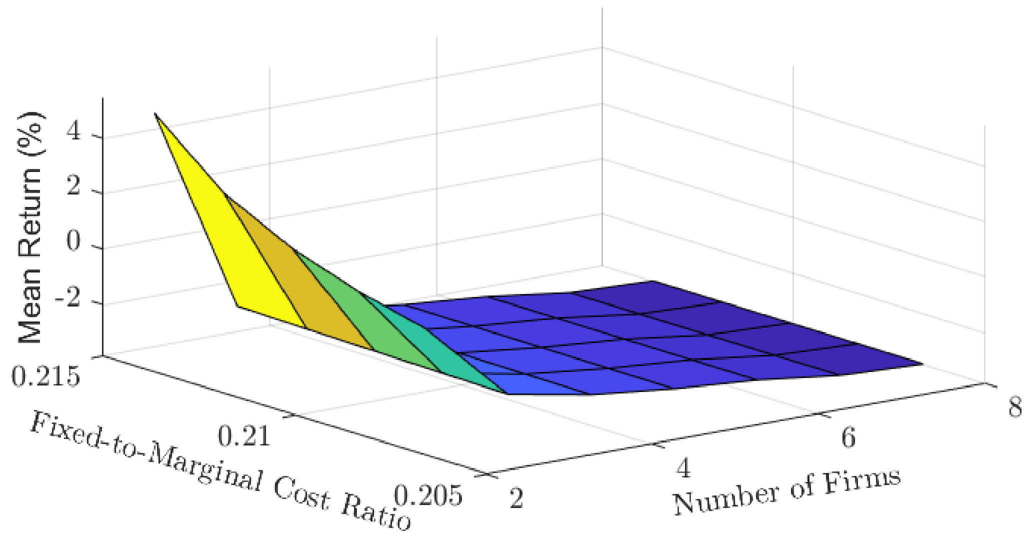
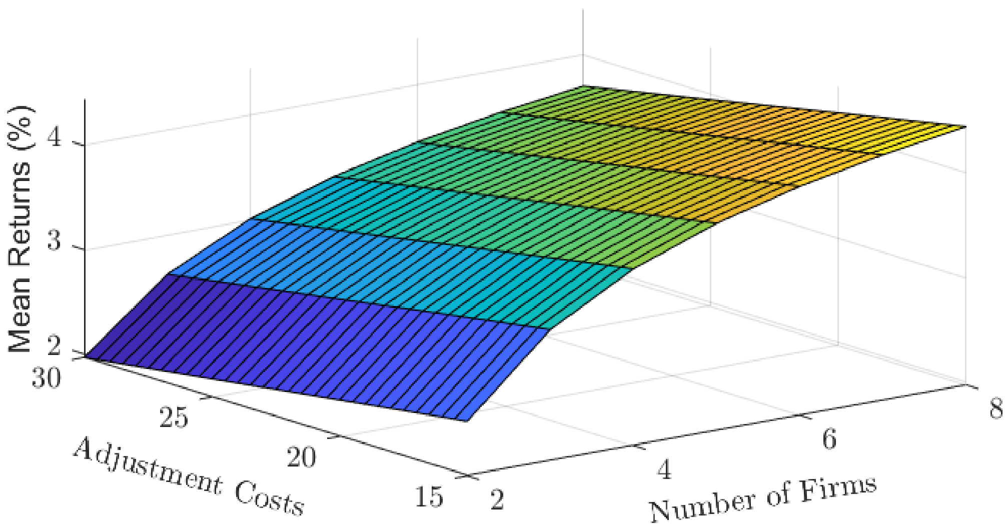


Figure 6. Industry concentration and equity returns with high adjustment costs

This figure displays, through a three-dimensional surface plot, the relation of mean equity returns and industry concentration (inversely related to the number of industry firms) for high levels of adjustment costs in the industry equilibrium analyzed in Section 5.1. The mean returns are computed from 1000 simulations of 59 yearly observations.



Capital Investment, Equity Returns and Aggregate Dynamics in Oligopolistic Production Economies

Online appendix

January 2024

A. Derivation of Optimal Consumption and Portfolio Policies

The representative consumer-investor's (CI's) optimization problem at any t is to

$$\max_{\mathbf{c}_t,\mathbf{q}_{t+1}} \mathcal{U}_t, \text{ s.t.,} \tag{A1.1}$$

$$\mathbf{p}_t \cdot \mathbf{c}_t \quad \leq \quad \mathbf{q}_t \cdot (\mathbf{d}_t + \mathbf{s}_t) - \mathbf{q}_{t+1} \cdot \mathbf{s}_t \equiv W_t. \tag{A1.2}$$

The Lagrangian with respect to (A1.1)-(A1.2) is

$$\max_{\mathbf{c}_t,\mathbf{q}_{t+1}} \mathcal{U}_t + \chi_t \left[W_t - \mathbf{p}_t \cdot \mathbf{c}_t \right], \; W_t = \mathbf{q}_t \cdot (\mathbf{d}_t + \mathbf{s}_t) - \mathbf{q}_{t+1} \cdot \mathbf{s}_t. \tag{A1.3}$$

Because preferences are strictly increasing in consumption, the budget constraint (A1.2) will be binding in optimum. Using concavity of the objective and convexity of the constraint, the optimal consumption and portfolio policies can be characterized in the standard fashion through a two-step process, where the optimal consumption vector $\mathbf{c}_t$ is first determined as a function of available consumption expenditure W_t, and the portfolio $\mathbf{q}_{t+1}$ is then determined taking as given the optimal consumption policy.

From the definition of the consumption basket $C_t \equiv \left[\sum_{j=1}^{J} \phi^j (c_{jt})^{(\sigma-1)/\sigma} \right]^{\sigma/(\sigma-1)}$, the first order optimality conditions for c_{jt} can be written

$$[(1-\alpha)(1-\eta)](C_t)^{\frac{1-\eta\sigma}{\sigma}} (c_{jt})^{-\frac{1}{\sigma}} \phi_j = \chi_t p_{jt}. \tag{A1.4}$$

Isolating c_t in (A1.4) and multiplying both sides by p_{jt} yields

$$p_{jt} c_{jt} = \chi_t^{-\sigma} (p_{jt})^{1-\sigma} (C_t)^{-(1-\eta\sigma)} (\phi_j)^{\sigma} [(1-\alpha)(1-\eta)]^{\sigma}. \tag{A1.5}$$

Then recognizing that $W_t = \sum_j p_{jt} c_{jt}$ and $P_t = \left[\sum_{j=1}^{J} (\phi_j)^{\sigma} (p_{jt})^{1-\sigma} \right]^{1/(1-\sigma)}$, summing both sides of (A1.5) over allows one to solve for the Lagrange multiplier as

$$\chi_t = \left(\frac{W_t}{P_t} \right)^{-\frac{1}{\sigma}} P_t^{-1} (C_t)^{\frac{1-\eta\sigma}{\sigma}} [(1-\alpha)(1-\eta)]. \tag{A1.6}$$

Substituting for χ_t in (A1.4) and rearranging terms then gives the optimal consumption functions

$$c_{jt}(\mathbf{p}_t^c, W_t) = \frac{W_t}{P_t} \left[\frac{P_t \phi_j}{p_{jt}} \right]^{\sigma}, \; j = 1, ..J. \tag{A1.7}$$

Now (A1.7) implies

$$(C_t)^{\frac{\sigma-1}{\sigma}} = \sum \phi_j (c_{jt})^{\frac{\sigma-1}{\sigma}} = \left(\frac{W_t}{P_t}\right)^{\frac{\sigma-1}{\sigma}} (P_t)^{\sigma-1} \left(\sum_j (\phi_j)^{\sigma} (p_{jt})^{1-\sigma}\right). \tag{A1.8}$$

But since $\sum_j (\phi_j)^{\sigma}(p_{jt})^{1-\sigma} = (P_t)^{1-\sigma}$, (A1.8) yields $C_t = \frac{W_t}{P_t}$. Next, conditional on optimal $\mathbf{c}_t$ (and hence $C_t = \frac{W_t}{P_t}$), the derivation of the optimal portfolio condition (6) in the text with Epstein and Zin (1989) preferences is standard using straightforward application of arguments in Epstein and Zin (1989). ∎

B1. Capacity utilization in Bertrand-Nash equilibrium

We start with the characterization of the punishment or Bertrand-Nash (BN) equilibrium path (where firms compete on prices) around its deterministic steady state. In the steady state of the BN equilibrium, the inverse demand, or pricing function in our model is

$$p(Y^{ind}) = M\phi(Y^{ind})^{-\frac{1}{\sigma}}, \tag{B1.1}$$

where $M = W^{\frac{1}{\sigma}} P^{\frac{(\sigma-1)}{\sigma}}$ and $Y^{ind} = \sum_{n=1}^{N} Y_{nj}$ is the industry output. Note that (B1.1) implies positive prices for any finite industry output. This leads to unique full capacity equilibrium. The proof is similar to Kreps and Scheinkman (1983) and hence the argument is only sketched here. Recall that the production function is $Y_t = A_t(u_t K_t)^{\psi}$. Since A_t and K_t are state variables (or given) in each period t, for notational ease we write $Y_t(u_t)$ as the firm-level output function of the capacity utilization choice u_t. Now fix any firm n. Suppose that the prices chosen by other firms, $i \neq n$, are $\underline{p}_{it} = p\left(\sum_{z=1}^{N} Y_t(u_{zt} = 1)\right)$. Then firm n clearly does not gain by choosing $p_{nt} < \underline{p}_{it}$ because, by construction, it can sell its full capacity at $\underline{p}_{nt}$. On the other hand, the firm will have zero revenues if it chooses $p_t > \underline{p}_{nt}$. Hence, $\tilde{p}_{nt} = p\left(\sum_{z=1}^{N} Y_t(u_{zt} = 1)\right) \equiv \underline{p}_t$ for all firms is a BN equilibrium price in the steady state. It follows that $\tilde{u}_t = 1$ is indeed the optimal capacity utilization strategy for each firm in this equilibrium. Suppose not. Any firm that chooses $u_{nt} < 1$ receives profits $\left(\underline{p}_t - h\right) Y_t(u_{nt}) < \left(\underline{p}_t - h\right) Y_t(\tilde{u}_t)$. We note that this argument applies for any distribution of production capacities $(K_{1t}, ...K_{Nt})$. Finally, the full-capacity BN equilibrium is clearly the worst equilibrium since it leads to the lowest feasible price. ∎

B2. Characterization of Bertrand-Nash equilibrium path

Put $y \equiv \frac{(\sigma(1-\psi)+\psi)}{\sigma}$. It is convenient to set,

$$\begin{aligned} \tilde{\Upsilon}_t &= M_t \phi N^{-\frac{1}{\sigma}} A_t^{\frac{(\sigma-1)}{\sigma}} K_t^{\frac{\psi(\sigma-1)}{\sigma}}, & \text{(B2.1)} \\ \tilde{\Upsilon}_t^K &= M_t \phi N^{-\frac{\sigma+1}{\sigma}} A_t^{\frac{\sigma-1}{\sigma}} K_t^{-y} \left(\frac{N\sigma - 1}{\sigma}\right). & \text{(B2.2)} \end{aligned}$$

Then,

$$\tilde{D}_t = \tilde{\Upsilon}_t - hA_tK_t^{\psi} - \Psi(\tilde{I}_t, K_t) - mK_t. \tag{B2.3}$$

We also recall, $\Psi(\tilde{I}_t, K_t) = \tilde{I}_t + 0.5\varphi\left(\frac{\tilde{I}_t}{K_t} - \delta\right)^2 K_t$. It is straightforward to compute:

$$\begin{aligned}
\Psi_I(\tilde{I}_t, K_t) &= 1 + \varphi\left(\frac{\tilde{I}_t}{K_t} - \delta\right), && \text{(B2.5)}\\
\Psi_K(\tilde{I}_t, K_t) &= -0.5\varphi\left[\left(\frac{\tilde{I}_{t+1}}{K_{t+1}}\right)^2 - \delta^2\right]. && \text{(B2.6)}
\end{aligned}$$

Now the optimal investment and firms' equity value along the BN equilibrium path satisfy:

$$\begin{aligned}
\frac{\partial \tilde{D}_t}{\partial I_t} &= \mathbb{E}_t\left[\frac{\partial \tilde{V}_{t+1}(K_{t+1})}{\partial I_{t+1}}\right], && \text{(B2.7)}\\
\tilde{V}_{t+1}(K_{t+1}) &= \mathbb{E}_t\left[\Omega_{t,t+1}\left(\tilde{D}_{t+1} + \tilde{V}_{t+2}\right)\right], && \text{(B2.8)}\\
K_{t+1} &= K_t(1-\delta) + I_t. && \text{(B2.9)}
\end{aligned}$$

Then, it follows from the foregoing that the Euler condition for investment satisfies

$$\begin{aligned}
0 \;=\; & -\left[1 + \varphi\left(\frac{\tilde{I}_t}{K_t} - \delta\right)\right] + \mathbb{E}_t\left[\Omega_{t,t+1}\left\{\frac{\partial \tilde{\Pi}_{t+1}}{\partial K_{t+1}} - m + \right.\right.\\
& \left.\left. 0.5\varphi\left[\left(\frac{\tilde{I}_{t+1}}{K_{t+1}}\right)^2 - \delta^2\right] + (1-\delta)\left(1 + \varphi\left(\frac{\tilde{I}_{t+1}}{K_{t+1}} - \delta\right)\right)\right\}\right], \qquad \text{(B2.10)}
\end{aligned}$$

where

$$\frac{\partial \tilde{\Pi}_{t+1}}{\partial K_{t+1}} = \psi\left[\tilde{\Upsilon}_t^K - hA_{t+1}K_{t+1}^{(\psi-1)}\right]. \tag{B2.11}$$

The resultant dividends are $\tilde{D}_t = \tilde{\Pi}_{t+1} - \Psi(\tilde{I}_t, K_t) - mK_t$ that determine cum- and ex-dividend equity values $\tilde{V}_t$, $\tilde{S}_t$ respectively through the asset market clearing condition $\mathbb{E}_t\left[\Omega_{t,t+1}\left(\frac{\tilde{D}_{t+1}+\tilde{S}_{t+1}}{\tilde{S}_t}\right)\right] = 1$. ∎

C1. Price defection and optimal capacity utilization ($\bar{u}_t^*$)

We argue that if any firm defects from $u_t^* < 1$, then it does so by producing at full capacity, that is, $\bar{u}_t^* = 1$. Suppose not and consider some candidate $\bar{u}_t^* < 1$. By the construction of the equilibrium, firms switch to the BN (punishment) equilibrium strategies following *any* defection from the tacit-collusive equilibrium path. Hence, the defecting firm's future payoffs subsequent to the defection are independent of $\bar{u}_t^*$. Therefore, it is the optimal strategy of the defecting firm to maximize the current operating profits (Π_t) from defection. We now prove that defection optimally involves choosing possibly a subsequence of lower prices arbitrarily close to p_t^* and selling output up to full capacity.

Note first that there exists some $\bar{N}_t(u_t^*)$ such that

$$A_t(K_t)^{\psi} \leq \bar{N}_t(u_t^*) A_t \left(u_t^* K_t\right)^{\psi}. \tag{C1.1}$$

In (C1.1), the left hand side (LHS) is the typical firm's full capacity output and $A_t \left(u_t^* K_t\right)^{\psi}$ in the RHS is the firm's output in the original equilibrium. Clearly, (C1.1) is satisfied if $\bar{N}_t(u_t^*) \geq (u_t^*)^{-\psi}$, where $\bar{N}_t(u_t^*)$ is well defined because $0 < u_t^* < 1$ and $0 < \psi < 1$. Then, let the defecting firm initially choose $p_t^* - \varepsilon$, for some small $\varepsilon > 0$; doing so allows it to sell at least the industry output in the original equilibrium, namely, $NA_t(u_t^* K_t)^{\psi}$. The residual capacity following this sale is no greater than max(0, $A_t(K_t)^{\psi} - NA_t(u_t^* K_t)^{\psi}$). The firm then chooses $p_t^* - 2\varepsilon$ and sells at least $NA_t(u_t^* K_t)^{\psi}$. After the second sale, max(0, $A_t(K_t)^{\psi} - 2NA_t(u_t^* K_t)^{\psi}$) is the upper bound on the residual capacity. The firm then charges $p_t^* - 3\varepsilon$ to sell at least $NA_t(u_t^* K_t)^{\psi}$ and so on. From the foregoing, the firm must be able to sell its full capacity through the sequence of prices $\left(p_t^* - \varepsilon, ..., p_t^* - \left[\frac{\bar{N}_t(u_t^*)}{N}\right]^{+} \varepsilon\right)$, where $\left[\frac{\bar{N}_t(u_t^*)}{N}\right]^{+}$ is the smallest integer greater than or equal to the ratio $\bar{N}_t(u_t^*)/N$. But since ε can be made arbitrarily small, it follows that the firm can sell its full capacity at price arbitrarily close to p_t^*. Note also that leaving residual capacity unsold following defection can not be optimal because the firm can sell that capacity at price close to $p_t^* > h$.

■

C2. Characterization of tacit-collusion equilibrium path

In the following, we suppress the '*' subscript for notational ease. Now, recall that the state along the equilibrium path is $\Gamma_t = (W_t, P_t, A_t, K_t)$. If there is a price defection at t, then firms next period play the BN (or punishment) equilibrium derived above, with the state $\Gamma_{t+1} = (W_{t+1}, P_{t+1}, A_{t+1}, K_{t+1})$, where $K_{t+1} = K_t(1-\delta) + I_t$. The BN value at $t+1$ will be denoted $\tilde{V}_{t+1}(K_{t+1})$, where using the analysis above (Section B.2)

$$\tilde{V}_{t+1}(K_{t+1}) = M_{t+1}\phi N^{-\frac{1}{\sigma}} A_{t+1}^{\frac{(\sigma-1)}{\sigma}} K_{t+1}^{\frac{\psi(\sigma-1)}{\sigma}} - hA_t K_{t+1}^{\psi} - \Psi(\tilde{I}_{t+1}, K_{t+1}) - mK_{t+1}, \tag{C2.1}$$

and $\tilde{I}_{t+1}$ is the optimal BN investment when the state is $\Gamma_{t+1} = (W_{t+1}, P_{t+1}, A_{t+1}, K_{t+1})$, that is, $\tilde{I}_{t+1} = K_{t+2} - (1-\delta)K_{t+1}$. Now, using the fact that along any candidate equilibrium path $\bar{u}_t = 1$ (established above), the Lagrangian for the constrained optimization problem with choice variables (u_t, I_t) is

$$\begin{aligned} \mathcal{L}_t \quad = \quad & D_t(u_t, I_t; K_t) + \mathbb{E}_t \left[V_{t+1}(K_{t+1})\right] + \varsigma_t \left(D_t(u_t, I_t; K_t) + \mathbb{E}_t \left[V_{t+1}(K_{t+1})\right] - \right. \\ & \left. \left\{\bar{D}_t(u_t, I_t; K_t) + \mathbb{E}_t \left[\tilde{V}_{t+1}(K_{t+1})\right]\right\}\right), \end{aligned} \tag{C2.2}$$

where $K_{t+1} = K_t(1-\delta) + I_t$ and

$$\begin{aligned} D_t(u_t, I_t; K_t) &= M_t \phi N^{-\frac{1}{\sigma}} A_{t+1}^{\frac{(\sigma-1)}{\sigma}} (u_t K_t)^{\frac{\psi(\sigma-1)}{\sigma}} - hA_t(u_t K_t)^{\psi} - \\ &\quad \Psi(I_t, K_t) - mK_t, \qquad \text{(C2.3)} \\ \bar{D}_t(u_t, I_t; K_t) &= M_t \phi N^{-\frac{1}{\sigma}} A_t^{\frac{(\sigma-1)}{\sigma}} K_t^{\frac{\psi(\sigma-1)}{\sigma}} u_t^{-\frac{\psi}{\sigma}} - hA_t K_t^{\psi} - \\ &\quad -\Psi(I_t, K_t) - mK_t. \qquad \text{(C2.4)} \end{aligned}$$

In (C2.4), we have used the fact that $\bar{u}_t = 1$. As above, we let $y \equiv \frac{\sigma(1-\psi)+\psi}{\sigma}$. It follows that (suppressing arguments of functions for notational convenience)

$$\begin{aligned} \frac{\partial D_t}{\partial u_t} &= \frac{\psi(\sigma-1)}{\sigma}\left[M_t \phi N^{-\frac{1}{\sigma}} A_t^{\frac{(\sigma-1)}{\sigma}} K_t^{\frac{\psi(\sigma-1)}{\sigma}} u_t^{-y}\right] - \psi h A_t K_t^{\psi} u_t^{\psi-1}, \qquad \text{(C2.5)} \\ \frac{\partial \bar{D}_t}{\partial u_t} &= -\frac{\psi}{\sigma}\left[M_t \phi N^{-\frac{1}{\sigma}} A_t^{\frac{(\sigma-1)}{\sigma}} K_t^{\frac{\psi(\sigma-1)}{\sigma}} u_t^{-\frac{(\psi+\sigma)}{\sigma}}\right], \qquad \text{(C2.6)} \\ \frac{\partial D_t^*}{\partial I_t} &= \frac{\partial \bar{D}_t}{\partial I_t} = -\Psi_I(I_t, K_t) = -\left[1 + \varphi\left(\frac{I_t}{K_t} - \delta\right)\right]. \qquad \text{(C2.7)} \end{aligned}$$

Now taking the first order conditions of $\mathcal{L}_t$ (in C2.2) with respect to (u_t, I_t, ς_t) yields the optimality conditions

$$\begin{aligned} \frac{\partial D_t}{\partial u_t}(1+\varsigma_t) &= \varsigma_t \frac{\partial \bar{D}_t}{\partial u_t}, \qquad \text{(C2.8)} \\ \left(-\left\{1 + \varphi\left(\frac{I_t}{K_t} - \delta\right)\right\} + \mathbb{E}_t\left[\frac{\partial V_{t+1}(K_{t+1})}{\partial I_t}\right]\right)(1+\varsigma_t) &= \varsigma_t\left(-\left\{1 + \varphi\left(\frac{I_t}{K_t} - \delta\right)\right\} + \right. \\ &\quad \left. \mathbb{E}_t\left[\frac{\partial \tilde{V}_{t+1}(K_{t+1})}{\partial I_t}\right]\right), \qquad \text{(C2.9)} \\ D_t + \mathbb{E}_t\left[V_{t+1}(K_{t+1})\right] &= \bar{D}_t + \mathbb{E}_t\left[\tilde{V}_{t+1}(K_{t+1})\right], \qquad \text{(C2.10)} \end{aligned}$$

where $K_{t+1} = K_t(1-\delta) + I_t$. Then eliminating ς_t in (C2.8)-(C2.9) by dividing both sides of (C2.9) by (C2.8) yields

$$\mathbb{E}_t\left[\frac{\partial V_{t+1}(K_{t+1})}{\partial I_t}\right] = \left[1 + \varphi\left(\frac{I_t}{K_t} - \delta\right)\right]\left(1 - \frac{\frac{\partial D_t}{\partial u_t}}{\frac{\partial \bar{D}_t}{\partial u_t}}\right) + \mathbb{E}_t\left[\frac{\partial \tilde{V}_{t+1}(K_{t+1})}{\partial I_t}\right]\frac{\frac{\partial D_t}{\partial u_t}}{\frac{\partial \bar{D}_t}{\partial u_t}}. \qquad \text{(C2.11)}$$

But from (C2.3)-C2.4), we get,

$$\chi_t \equiv \frac{\left(\frac{\partial D_t}{\partial u_t}\right)}{\left(\frac{\partial \bar{D}_t}{\partial u_t}\right)} = \left(\frac{u_t^{\psi}}{p_t}\right)(p_t - \sigma(p_t - h)) \tag{C2.12}$$

$$= -(\sigma-1)u_t^{\psi} + \frac{\sigma h N^{\frac{1}{\sigma}} A_t^{\frac{1}{\sigma}} K_t^{\frac{\psi}{\sigma}} u_t^{\frac{\psi(\sigma+1)}{\sigma}}}{M_t \phi}. \tag{C2.13}$$

Hence (C2.9) can be written as

$$\mathbb{E}_t\left[\frac{\partial V_{t+1}(K_{t+1})}{\partial I_t}\right] = \left[1+\varphi\left(\frac{I_t}{K_t}-\delta\right)\right](1-\chi_t) + \mathbb{E}_t\left[\frac{\partial \tilde{V}_{t+1}(K_{t+1})}{\partial I_t}\right]\chi_t. \tag{C2.14}$$

Furthermore, using standard techniques we have

$$\mathbb{E}_t\left[\frac{\partial V_{t+1}}{\partial I_t}\right] = \mathbb{E}_t\Bigg[\Omega_{t,t+1}\Bigg\{\frac{\partial \Pi_{t+1}(u_{t+1}K_{t+1})}{\partial K_{t+1}} - m + 0.5\varphi\left(\left(\frac{I_{t+1}}{K_{t+1}}\right)^2 - \delta^2\right) + (1-\delta)\mathbb{E}_{t+1}\left[\frac{\partial V_{t+2}}{\partial K_{t+2}}\right]\Bigg\}\Bigg]. \tag{C2.15}$$

But $\mathbb{E}_{t+1}\left[\frac{\partial V_{t+2}(K_{t+2})}{\partial I_{t+1}}\right] = \mathbb{E}_{t+1}\left[\frac{\partial V_{t+2}}{\partial K_{t+2}}\right]$ and hence forward induction on (C2.15) yields

$$\mathbb{E}_{t+1}\left[\frac{\partial V_{t+2}}{\partial K_{t+2}}\right] = \left[1+\varphi\left(\frac{I_{t+1}}{K_{t+1}}-\delta\right)\right](1-\chi_{t+1}) + \mathbb{E}_{t+1}\left[\frac{\partial \tilde{V}_{t+2}(K_{t+2})}{\partial I_{t+1}}\right]\chi_{t+1}, \tag{C2.16}$$

where χ_{t+1} is defined as in (C2.12)-(C2.13) for $t+1$. However, since $\mathbb{E}_{t+1}\left[\frac{\partial \tilde{V}_{t+2}(K_{t+2})}{\partial I_{t+1}}\right] = \mathbb{E}_{t+1}\left[\frac{\partial V_{t+2}}{\partial K_{t+2}}\right]$, it follows from investment optimality along the punishment equilibrium path (see Section B.2)

$$\mathbb{E}_{t+1}\left[\frac{\partial \tilde{V}_{t+2}(K_{t+2})}{\partial K_{t+2}}\right] = 1+\varphi\left(\frac{\tilde{I}_{t+1}}{K_{t+1}}-\delta\right). \tag{C2.17}$$

Then substituting (C2.17) in (C2.16) and returning to (C2.15), we get

$$\mathbb{E}_t\left[\frac{\partial V_{t+1}}{\partial I_t}\right] = \mathbb{E}_t\Bigg[\Omega_{t,t+1}\Bigg\{\frac{\partial \Pi_{t+1}(K_{t+1})}{\partial K_{t+1}} - m + 0.5\varphi\left(\left(\frac{I_{t+1}}{K_{t+1}}\right)^2 - \delta^2\right) + (1-\delta)\times \left(1+\varphi\left(\left[\frac{I_{t+1}(1-\chi_{t+1}) + \chi_{t+1}\tilde{I}_{t+1})}{K_{t+1}}\right] - \delta\right)\right)\Bigg\}\Bigg], \tag{C2.18}$$

where

$$\Pi_{t+1}(u_{t+1},K_{t+1}) = M_{t+1}\phi N^{-\frac{1}{\sigma}} A_{t+1}^{\frac{(\sigma-1)}{\sigma}}(u_{t+1}K_{t+1})^{\frac{\psi(\sigma-1)}{\sigma}} - hA_{t+1}(u_{t+1}K_{t+1})^{\psi}, \quad (C2.19)$$

$$\frac{\partial \Pi_{t+1}}{\partial K_{t+1}} = \psi\left[\frac{(\sigma-1)}{\sigma}M_{t+1}\phi N^{-\frac{1}{\sigma}} A_{t+1}^{\frac{(\sigma-1)}{\sigma}} u_{t+1}^{\frac{\psi(\sigma-1)}{\sigma}} K_{t+1}^{-y} - hA_{t+1}u_{t+1}^{\psi}K_{t+1}^{\psi-1}\right]. \quad (C2.20)$$

Furthermore, along the BN equilibrium path at t,

$$\mathbb{E}_t\left[\frac{\partial \tilde{V}_{t+1}(K_{t+1})}{\partial I_t}\right] = \mathbb{E}_t\left[\Omega_{t,t+1}\left\{\frac{\partial \tilde{\Pi}_{t+1}(K_{t+1})}{\partial K_{t+1}} - m + 0.5\varphi\left(\frac{\tilde{I}_{t+1}}{K_{t+1}} - \delta\right)^2 + (1-\delta)\times \left(1+\varphi\left(\frac{\tilde{I}_{t+1}}{K_{t+1}} - \delta\right)\right)\right\}\right], \quad (C2.21)$$

where

$$\tilde{\Pi}_{t+1}(K_{t+1}) = M_{t+1}\phi N^{-\frac{1}{\sigma}} A_{t+1}^{\frac{(\sigma-1)}{\sigma}} K_{t+1}^{\frac{\psi(\sigma-1)}{\sigma}} - hA_{t+1}K_{t+1}^{\psi} \quad (C2.22)$$

$$\frac{\partial \tilde{\Pi}_{t+1}(K_{t+1})}{\partial K_{t+1}} = \psi\left[\frac{(\sigma-1)}{\sigma}M_{t+1}\phi N^{-\frac{1}{\sigma}} A_{t+1}^{\frac{(\sigma-1)}{\sigma}} K_{t+1}^{-y} - hA_{t+1}K_{t+1}^{\psi-1}\right] \quad (C2.23)$$

Hence it follows from (C2.20) and (C2.23) that

$$\frac{\partial \Pi_{t+1}}{\partial K_{t+1}} = u_{t+1}^{\frac{\psi(\sigma-1)}{\sigma}}\frac{\partial \tilde{\Pi}_{t+1}}{\partial K_{t+1}} - \psi hA_{t+1}K_{t+1}^{\psi-1}u_{t+1}^{\psi}[1-u_{t+1}^{-\frac{\psi}{\sigma}}] \quad (C2.24)$$

Hence,

$$\frac{\partial \tilde{\Pi}_{t+1}}{\partial K_{t+1}} = (u_{t+1})^{-\frac{\psi(\sigma-1)}{\sigma}}\frac{\partial \Pi_{t+1}}{\partial K_{t+1}} + \psi hA_{t+1}K_{t+1}^{\psi-1}(u_{t+1})^{\frac{\psi}{\sigma}}[1-u_{t+1}^{-\frac{\psi}{\sigma}}] \quad (C2.25)$$

$$\frac{\partial \Pi_{t+1}}{\partial K_{t+1}} - \chi_t\frac{\partial \tilde{\Pi}_{t+1}}{\partial K_{t+1}} = \frac{\partial \Pi_{t+1}}{\partial K_{t+1}}(1-\chi_t u_{t+1}^{-\frac{\psi(\sigma-1)}{\sigma}}) - \chi_t\psi hA_{t+1}K_{t+1}^{(\psi-1)}[u_{t+1}^{\frac{\psi}{\sigma}} - 1] \quad (C2.26)$$

Then using (C2.18) and (C2.26) in (C2.14) and rearranging terms yields the optimality condition:

$$\left[1+\varphi\left(\frac{I_t}{K_t}-\delta\right)\right](1-\chi_t) =$$

$$\mathbb{E}_t\left[\Omega_{t,t+1}\left\{\frac{\partial \Pi_{t+1}}{\partial K_{t+1}}\left(1-\chi_t u_{t+1}^{-\frac{\psi(\sigma-1)}{\sigma}}\right) - \chi_t\psi[u_{t+1}^{\frac{\psi}{\sigma}}-1]hA_{t+1}K_{t+1}^{\psi-1} - (1-\chi_t)m + 0.5\varphi\left[\left(\frac{I_{t+1}}{K_{t+1}}\right)^2 - \chi_t\left(\frac{\tilde{I}_{t+1}}{K_{t+1}}\right)^2 - \delta^2(1-\chi_t)\right] + (1-\delta)\left[(1-\chi_t)+\varphi\left(\frac{I_{t+1}(1-\chi_{t+1})-\tilde{I}_{t+1}(\chi_t-\chi_{t+1})}{K_{t+1}} - (1-\chi_t)\delta\right)\right]\right\}\right]. \quad (C2.27)$$

■

C3. Non-viability of tacit collusion for large number of firms

Utilizing the notation in (beginning of) Section 3, we denote by $\hat{V}_t(\Gamma_t)$ the equity value of a monopoly in the typical sector, condition on the state Γ_t. Along a symmetric equilibrium with binding TCIC, and N firms in the sector, the per-firm equity value with state Γ_t is $V_t^*(\Gamma_t; N) < \frac{\hat{V}_t(\Gamma_t)}{N}$. Then fix any subgame Γ_t and hence $\hat{V}_t(\Gamma_t)$. Because $\frac{\hat{V}_t(\Gamma_t)}{N}$ is a continuous and strictly decreasing function of N that converges to zero as $N \to \infty$, it follows that for every $\epsilon > 0$, there exists some $N_\epsilon(\Gamma_t)$ such that $V_t^*(\Gamma_t; N) < \frac{\hat{V}_t(\Gamma_t)}{N} < \epsilon$ if $N \geq N_\epsilon(\Gamma_t)$. Since this argument applies for any t and Γ_t, for every $\epsilon > 0$, we can set $\bar{N}_\epsilon(\Gamma_t) = \max_{\Gamma_{t+1}(\Gamma_t)} \{N_\epsilon(\Gamma_{t+1})\}$, where $\Gamma_{t+1}(\Gamma_t)$ denotes subgames Γ_{t+1} that lie on the continuation game from Γ_t.

Meanwhile, for every Γ_t, the BN equity value $\tilde{V}_t(\Gamma_t; N) \geq 0$. It follows that $(V_t^*(\Gamma_t; N) - \tilde{V}_t(\Gamma_t; N)) < \epsilon$ if $N \geq N_\epsilon$. Now the markup $(p_t^*(\Gamma_t; N) - h)$ along the tacit-collusive path stays strictly bounded away from zero, in contrast to the BN equilibrium. Since $u_t^*(\Gamma_t; N) < 1$ in a binding TCIC (as shown in the text), it follows from the foregoing that Γ_t, there exists some $\bar{N}(\Gamma_t)$ such that

$$(p_t^*(\Gamma_t; N) - h)A_t(K_t(\Gamma_t; N))^\psi(1 - (u_t^*(\Gamma_t; N))^\psi) > \mathbb{E}_t\left[V_{t+1}^*(\Gamma_{t+1}) - \tilde{V}_{t+1}(\Gamma_{t+1})\right], \qquad \text{(C3.1)}$$

for $N \geq \bar{N}(\Gamma_t)$, which violates the TCIC. ∎

C4. Existence of deterministic steady state

In the deterministic steady state, sectoral productivity A is time-invariant and hence the endogenous variables are also time-invariant. In the steady state of the BN equilibrium, the industry price is set when production is at full capacity, that is, $\tilde{u} = 1$, following the arguments given above (Section B.1). The equilibrium capital stock is $\tilde{K}$ and hence $\tilde{I} = \delta\tilde{K}$, which satisfies the capital transition condition $\tilde{K} = \tilde{K}(1-\delta) + \tilde{I}$. $\tilde{K}$ is determined under the assumption that the adjustment costs in the transition from K to $\tilde{K}$ following a defection are amortized over the infinite horizon, so that the per-period adjustment costs are zero in the punishment equilibrium.[1] It follows that $\Psi(\tilde{I}, \tilde{K}) = \tilde{I} = \delta\tilde{K}$ (since $\left(\frac{\tilde{I}}{\tilde{K}} - \delta\right) = 0$).

Now, suppose that the initial symmetric capital stock distribution is $\tilde{K}$ for all firms. For any firm n, let $\tilde{K}_{-n}$ denote the capital stocks of each of the other firms. Then (with $\tilde{u}_n = 1$), the period operating profits are

$$\Pi(\tilde{K}_{-n}, \tilde{K}) = \left(M\phi A^{-\frac{1}{\sigma}}[(N-1)\tilde{K}^\psi_{-n} + \tilde{K}^\psi]^{-\frac{1}{\sigma}} - h\right) A\tilde{K}^\psi, \qquad \text{(C4.1)}$$

[1] Along the equilibrium path, once industry firms switch to the punishment equillibrium path following defection, then they adopt the optimal punishment path investment policies specified in equation (B2.10) above, which involves adjustment costs. However, defection in the steady state involves transitioning from the steady state capital stock K to $\tilde{K}$. Since the deterministic steady state holds to perpetuity, amortization of the firxed adjustment costs of the transition yields per period adjustment costs of zero, consistent with the constrained maximization problem set up in (C4.20)-(C4.21).

and dividends are

$$D(I, \tilde{K}_{\_n}, \tilde{K}) = \Pi(\tilde{K}_{-n}, \tilde{K}) - \Psi(I, \tilde{K}) - m\tilde{K}. \tag{C4.2}$$

Furthermore, the firm's steady state value function satisfies the optimization problem:

$$V(\tilde{K}_{-n}, \tilde{K}) = \max_{I} D(I, \tilde{K}_{-n}, \tilde{K}) + \alpha V(\tilde{K}((1-\delta) + I). \tag{C4.3}$$

$\tilde{K}$ is then steady state equilibrium for all firms if the solution to (C4.3) implies $\tilde{I} = \delta\tilde{K}$. Hence, $\Psi(\tilde{I}, \tilde{K}) = \delta\tilde{K}$ (since $\left(\frac{I}{K} - \delta\right) = 0$). The Euler condition for (C4.3) evaluated at $\tilde{I} = \delta\tilde{K}$ is:

$$1 = \alpha \left[\frac{\partial \tilde{\Pi}}{\partial K} - m + (1-\delta) \right], \tag{C4.4}$$

where for notational ease $\frac{\partial \Pi}{\partial K} = \frac{\partial \Pi(\tilde{K}_{-n}, K)}{\partial K} |_{K=\tilde{K}}$. To compute this, we rewrite (C4.1) above as

$$\Pi(\tilde{K}_{-n}, K) = M\phi A^{\frac{\sigma-1}{\sigma}}((N-1)\tilde{K}_{-n}^{\psi} + K^{\psi})^{-\frac{1}{\sigma}} K^{\psi} - hAK^{\psi}. \tag{C4.5}$$

Hence, defining $y \equiv \frac{(\sigma(1-\psi)+\psi)}{\sigma}$,

$$\frac{\partial \Pi(\tilde{K}_{-n}, K)}{\partial K} |_{K=\tilde{K}} = \psi M\phi A^{\frac{\sigma-1}{\sigma}} N^{-\frac{\sigma+1}{\sigma}} \tilde{K}^{-y} \left[-\frac{1}{\sigma} + N \right] - \psi hA\tilde{K}^{(\psi-1)}, \tag{C4.6}$$

which can be written as

$$\frac{\partial \tilde{\Pi}}{\partial K} = \psi \left[\psi M\phi A^{\frac{\sigma-1}{\sigma}} N^{-\frac{\sigma+1}{\sigma}} \tilde{K}^{-y} \left(\frac{N\sigma - 1}{\sigma} \right) - hA\tilde{K}^{(\psi-1)} \right]. \tag{C4.7}$$

For convenience, we will set

$$\tilde{\Upsilon} \equiv M\phi N^{-\frac{1}{\sigma}} \tilde{K}^{\frac{\psi(\sigma-1)}{\sigma}} A^{\frac{(\sigma-1)}{\sigma}}, \tag{C4.8}$$

$$\tilde{\Upsilon}^{K} \equiv M\phi A^{\frac{\sigma-1}{\sigma}} N^{-\frac{\sigma+1}{\sigma}} \tilde{K}^{-y} \left(\frac{N\sigma - 1}{\sigma} \right), \tag{C4.9}$$

so that

$$\frac{\partial \tilde{\Pi}}{\partial K} = \psi \left[\tilde{\Upsilon}^{K} - hA\tilde{K}^{(\psi-1)} \right]. \tag{C4.10}$$

Then the Euler condition (C4.4) and dividends $\tilde{D}$ are given by

$$1 = \alpha \left[\psi \left(\tilde{\Upsilon}^{K} - hA\tilde{K}^{(\psi-1)} \right) - m + (1-\delta) \right], \tag{C4.11}$$

$$\tilde{D} = \tilde{\Upsilon} - hA\tilde{K}^{\psi} - (\delta + m)\tilde{K}. \tag{C4.12}$$

Now, for every given (W, P) and hence M $(= W^{\frac{1}{\sigma}}P^{\frac{(\sigma-1)}{\sigma}})$, (C4.9) and (C4.11) imply that $\frac{\partial^2\tilde{\Pi}}{\partial K^2} < 0$ (since $y > 0$ and $\psi < 1$). Furthermore, $\lim_{K\to 0}\frac{\partial\tilde{\Pi}}{\partial K} = \infty$ and $\lim_{K\to\infty}\frac{\partial\tilde{\Pi}}{\partial K} = 0$. Hence, by continuity of $\tilde{\Pi} = \tilde{\Upsilon} - hA\tilde{K}^{\psi}$ in K, it follows that there exists a unique optimal $\tilde{K}$, namely, the solution to (C4.11). This implies from (C4.12) that $\tilde{D}$ is well defined as is $\tilde{p} = M\phi(NA\tilde{K}^{\psi})^{-1/\sigma}$. Hence, $\tilde{W} = \sum_j \tilde{D}_j + 1 - \alpha$ (where $1-\alpha$ is the net income from the risk free asset). Hence, $\tilde{W}$ is well-defined, as is $\tilde{P}$. It also follows that cum- and ex-dividend equity values are well defined in the steady state of the punishment equilibrium since $\tilde{V} = \frac{\tilde{D}}{1-\alpha}$, $\tilde{S} = \frac{\alpha\tilde{D}}{1-\alpha}$.

We turn now to the definition and existence of the steady state for the (tacit-collusive) equilibrium path. As above, we suppress the '*' subscript for notational ease. The firm's equilibrium choices are represented by (u, I, K), with $I = \delta K$. Hence, $\Psi(I, K) = I$. For further analysis, it is useful to define

$$\Upsilon \equiv M\phi N^{-\frac{1}{\sigma}}K^{\frac{\psi(\sigma-1)}{\sigma}}A^{\frac{(\sigma-1)}{\sigma}}u^{-\frac{\psi}{\sigma}}, \tag{C4.13}$$

$$\Upsilon^u \equiv M\phi N^{-\frac{1}{\sigma}}(uK)^{\frac{\psi(\sigma-1)}{\sigma}}A^{\frac{(\sigma-1)}{\sigma}}, \tag{C4.14}$$

$$\Upsilon^K \equiv \frac{\psi(\sigma-1)\Upsilon^u}{\sigma K}. \tag{C4.15}$$

The, for any (u, I), the per-period dividends are:

$$D(u, I; K) = \Upsilon^u - hA\,(uK)^{\psi} - I - mK. \tag{C4.16}$$

Using Section B.2, the optimal price defection sets $\bar{u} = 1$. Hence, the defecting dividends can be written

$$\bar{D}(u, I; K) = \Upsilon - hAK^{\psi} - I - mK. \tag{C4.17}$$

Therefore, $\bar{D}(u, I; K) - D(u, I; K)$ equals

$$\begin{aligned} &\left[M\phi N^{-\frac{1}{\sigma}}(uK)^{-\frac{\psi}{\sigma}}A^{-\frac{1}{\sigma}} - h\right]AK^{\psi}(1-u^{\psi}) \\ =\; &\left[\Upsilon - hAK^{\psi}\right](1-u^{\psi}). \end{aligned} \tag{C4.18}$$

It follows then that

$$\begin{aligned} \frac{\partial D(u, I; K)}{\partial u} &= \left(\frac{\psi}{u}\right)\left[\frac{(\sigma-1)}{\sigma}\Upsilon^u - hA\,(uK)^{\psi}\right], \\ \frac{\partial \bar{D}(u, I; K)}{\partial u} &= -\left(\frac{\psi}{u\sigma}\right)\Upsilon. \end{aligned} \tag{C4.19}$$

$$\frac{\frac{\partial D}{\partial u}}{\frac{\partial \bar{D}}{\partial u}} = \left(\frac{u^{\psi}}{p}\right)(p - \sigma(p-h))$$
$$= -(\sigma-1)u^{\psi} + \frac{\sigma h (AN)^{\frac{1}{\sigma}} K^{\frac{\psi}{\sigma}} u^{\frac{\psi(\sigma+1)}{\sigma}}}{M\phi} \equiv \chi. \tag{C4.20}$$

Then the deterministic steady state of the equilibrium is represented by the time-invariant version of the tacit-collusion Euler condition (C2.27) under the assumption that the adjustment costs in the transition from K to $\tilde{K}$ following a defection are amortized over the infinite horizon, so that the per-period adjustment costs are zero. Hence, in the steady state:

1. (u, K) satisfy the condition

$$(1-\chi) = \alpha \left[\frac{\partial \Pi}{\partial K}(1 - \chi u^{-\frac{\psi(\sigma-1)}{\sigma}}) - \chi \psi u^{\frac{\psi}{\sigma}}[1 - u^{-\frac{\psi}{\sigma}}]hAK^{(\psi-1)} + (1-\chi)((1-\delta)-m)\right]. \tag{C4.21}$$

   Equilibrium and defecting dividends are given by

$$D = \Upsilon^{u} - hA\,(uK)^{\psi} - (m+\delta)K, \tag{C4.22}$$
$$\bar{D} = \Upsilon - hAK^{\psi} - (m+\delta)K. \tag{C4.23}$$

   The tacit-collusion incentive compatibility condition (TCIC) is satisfied:

$$\left(\bar{D} - D\right) = \frac{\alpha}{1-\alpha}\left(D - \tilde{D}\right), \tag{C4.24}$$

   where $\tilde{D}$ is given by (C4.12).

2. Product price in the representative sector is

$$p = M\phi A^{-\frac{1}{\sigma}} N^{-\frac{1}{\sigma}} (uK)^{-\frac{\psi}{\sigma}}. \tag{C4.25}$$

   Aggregate income and price index is given by

$$W = \sum_{j} D_j + 1 - \alpha, \tag{C4.26}$$
$$P = \left[\sum_{j=1}^{J} \phi_j^{\sigma} p_j^{1-\sigma}\right]^{1/(1-\sigma)}. \tag{C4.27}$$

The analysis for the existence of solutions to (C4.21) proceeds as follows. Let,

$$\begin{aligned} \upsilon(u,K) &= -(1-\chi)+\alpha\left[\frac{\partial\Pi}{\partial K}(1-\chi u^{-\frac{\psi(\sigma-1)}{\sigma}})-\chi\psi u^{\frac{\psi}{\sigma}}[1-u^{-\frac{\psi}{\sigma}}]hAK^{(\psi-1)}+\right.\\ &\left.(1-\chi)((1-\delta)-m)\right]. \end{aligned} \tag{C4.28}$$

Note that $\upsilon(u,K)$ is continuous on its domain $[0,1]\times\mathcal{R}_{+}$. Furthermore, we have

$$\frac{\partial\Pi}{\partial K}=\Upsilon^{K}-\psi hAu^{\psi}K^{(\psi-1)}. \tag{C4.29}$$

It is then straightforward to check from the foregoing that for any $1\geq u>0$

$$\lim_{K\to0}\chi=0,\ \lim_{K\to\infty}\chi=\infty,\ \lim_{K\to0}\frac{\partial\Pi}{\partial K}=\infty,\ \lim_{K\to\infty}\frac{\partial\Pi}{\partial K}=0. \tag{C4.30}$$

Furthermore, from (C4.20) and (C4.29), if $\left(\frac{\psi}{\sigma}+\psi-1\right)<0$, that is, if

$$\frac{\sigma}{\sigma+1}>\psi, \tag{C4.31}$$

then it can be shown that

$$\lim_{K\to0}\upsilon(u,K)=\infty,\ \lim_{K\to\infty}\upsilon(u,K)=\infty. \tag{C4.32}$$

Hence, from continuity, (C4.32) implies that for every $1\geq u>0$, there exist well defined zeros of the steady state Euler condition (C4.21). The local optimality of these zeros can be checked by standard methods since $\upsilon(u,K)$ is also differentiable. ∎

D. Derivation of Equations in Section 5

Let $\bar{Y}_t\equiv A_t(K_t)^{\psi}$ and $Y_t\equiv A_t(u_tK_t)^{\psi}$. Using Equation (12) in the text, the (binding) TCIC is thus

$$(p_t-h)(\bar{Y}_t-Y_t)-\mathbb{E}_t\left[V_{t+1}(\Gamma_{t+1})-\tilde{V}_{t+1}(\Gamma_{t+1})\right]=0. \tag{D1.1}$$

Now consider the case where N_t is increased to N_{t+1}, while keeping fixed the aggregate quantities W_t and P_t. We will denote the equilibrium path with the additional active firm by subscript "+." The TCIC is then

$$(p_t^{+}-h)(\bar{Y}_t^{+}-Y_t^{+})-\mathbb{E}_t\left[V_{t+1}^{+}(\Gamma_{t+1})-\tilde{V}_{t+1}^{+}(\Gamma_{t+1})\right]=0. \tag{D1.2}$$

Subtracting (D1.2) from (D1.1) gives

$$\Delta_N(p_t) = \frac{\Delta_{N_t}\left(\mathbb{E}_t\left[V_{t+1}(\tilde{\Gamma}_{t+1}) - \bar{V}_{t+1}(\tilde{\Gamma}_{t+1})\right]\right) - (p_t - h)\Delta_{N_t}(\bar{Y}_t - Y_t)}{(\bar{Y}_t - Y_t)}. \tag{D1.3}$$

Note that $\bar{Y}_t - Y_t > 0$ because $\bar{u}_t = 1 > u_t$. Hence, the sign of $\Delta_{N_t}(p_t)$ is determined by the punishment and market dilution effects defined in Equation (18).

Next, we denote by $R_{t+1}(N)$ and $R_{t+1}(N+1)$ the gross $t+1-$period return of with N and $(N+1)$ firms, respectively, holding other things fixed. Furthermore, for notational simplicity, we suppress the dependence on the "status quo" number of firms N. Then,

$\Delta_N\left(Cov_t\left(\Omega_{t,t+1}, R_{t+1}\right) R_{f,t+1}\right) =$

$$\begin{aligned}
& [Cov_t\left(\Omega_{t,t+1}(N+1), R_{t+1}(N+1)\right) - Cov_t(\Omega_{t,t+1}, R_{t+1})](R_{f,t+1} + \Delta_N(R_{f,t+1})) \\
\simeq\ & [Cov_t\left(\Omega_{t,t+1}, R_{t+1}(N+1) - R_{t+1}\right) + Cov_t\left(\Omega_{t,t+1}(N+1) - \Omega_{t,t+1}, R_{t+1}(N+1)\right)]R_{f,t+1} \\
\simeq\ & [Cov_t(\Omega_{t,t+1}, \Delta_N(R_{t+1})) + Cov_t(\Delta_N(\Omega_{t,t+1}), R_{t+1}(N+1))]R_{f,t+1} \\
\simeq\ & Cov_t\left(\Omega_{t,t+1}, \Delta_N\left(\frac{D_{t+1} + S_{t+1}}{S_t}\right)\right) R_{f,t+1},
\end{aligned} \tag{D1.4}$$

when $\Delta_N(\Omega_{t,t+1})$ and $\Delta_N(R_{f,t+1})$ are small. ■

E. Loglinear approximations (Section 5.1)

E1. Stochastic discount factor

Since $C_t = \frac{W_t}{P_t}$, the SDF in Equation (4) in the text can be written (see Epstein and Zin 1989)

$$\frac{\Lambda_{t+1}}{\Lambda_t} = \alpha^\theta \left(\frac{G^W_{t+1}}{G^P_{t+1}}\right)^{-\eta\theta} R^{\theta-1}_{C,t+1}, \theta \equiv \frac{1-\gamma}{1-\eta}, \tag{E1.0}$$

where G^W_{t+1} and G^P_{t+1} are the gross growth rates of W and P, respectively, and $R_{C,t+1}$ is the gross return on the asset that pays out aggregate consumption. Therefore, from Equation (E1.0), the log of the real pricing kernel $\lambda_{t+1} \equiv \log(\Lambda_{t,t+1}/\Lambda_t)$ is

$$\lambda_{t+1} = \theta \log \alpha - \eta\theta \left[g_{w,t+1} - g_{\pi,t+1}\right] + (\theta - 1) r_{c,t+1}, \tag{E1.1}$$

where $g_{w,t+1}$ and $g_{\pi,t+1}$ are the log growth rates of income and aggregate price index, respectively, and $r_{c,t+1} \equiv \log(R_{C,t+1})$. Using the Campbell and Shiller (1988) log-linearization, we can represent $r_{c,t+1}$ as

$$r_{c,t+1} = \varkappa_{c0} + \varkappa_{c1}\ell_{c,t+1} - \ell_{ct} + g_{w,t+1} - g_{\pi,t+1}, \tag{E1.2}$$

where ℓ_{ct} is the log price-consumption ratio. Here

$$\varkappa_{c0} = \log(1 + \exp(\ell_c)) - \varkappa_{c1}\ell_c; \varkappa_{c1} = \frac{\exp(\ell_c)}{1 + \exp(\ell_c)}, \tag{E1.3}$$

where ℓ_c is the unconditional mean of ℓ_{ct}. We guess and then verify that ℓ_{ct} is a linear function of the logs of the aggregate state variables W_t and P_t, that is,

$$\ell_{ct} = \kappa_{c0} + \kappa_{cw} w_t + \kappa_{c\pi}\pi_t. \tag{E1.4}$$

Now, $g_{w,t+1} = (\rho_w - 1)w_t + \varepsilon^w_{t+1}$ and $g_{\pi,t+1} = (\rho_\pi - 1)\pi_t + \varepsilon^\pi_{t+1}$. Then substitution of (E1.3) and (E1.4) in (E1.2) gives

$$\lambda_{t+1} = B_0 + B_w w_t + B_\pi \pi_t + b_w \varepsilon_{w,t+1} + b_\pi \varepsilon_{\pi,t+1}, \tag{E1.5}$$

where

$$\begin{aligned}
B_0 &\equiv \theta \log \alpha; \\
B_w &\equiv (\rho_w - 1)\left[\theta(1-\eta) - 1\right] + (\theta - 1)\kappa_{cw}[\varkappa_{c1}\rho_w - 1]; \\
B_\pi &\equiv (\rho_\pi - 1)\left[\theta(\eta - 1) + 1\right] + (\theta - 1)\kappa_{c\pi}[\varkappa_{c1}\rho_\pi - 1]; \\
b_w &\equiv -\eta\theta + (\theta - 1)[\varkappa_{c1}\kappa_{cw} + 1]; \\
b_\pi &\equiv \eta\theta - (\theta - 1)[1 - \varkappa_{c1}\kappa_{c\pi}].
\end{aligned} \tag{E1.6}$$

We can then obtain the coefficients of ℓ_{ct} in (E1.4) through the method of undetermined coefficients. The Euler condition for returns yields

$$\begin{aligned}
1 &= \mathbb{E}_t[\exp(\lambda_{t+1} + \pi_t - \pi_{t+1} + r_{c,t+1})] \\
&= \mathbb{E}_t[\exp(\theta \log \alpha - \eta\theta\left[g_{w,t+1} - g_{\pi,t+1}\right] + \theta r_{c,t+1} + \pi_t - \pi_{t+1})]
\end{aligned} \tag{E1.7}$$

Since (E1.7) must hold for all values of the state variables, all terms involving w_t and π_t must satisfy

$$w_t\theta[(\rho_w - 1)(1 - \eta) + \kappa_{cw}(\varkappa_{c1}\rho_w - 1)] = 0, \tag{E1.8}$$

$$\pi_t\left\{\theta[(\rho_\pi - 1)(\eta - 1) + \kappa_{c\pi}(\varkappa_{c1}\rho_\pi - 1)] + 1 - \rho_\pi\right\} = 0. \tag{E1.9}$$

From (E1.8)-(E1.9), it follows

$$\kappa_{cw} = \frac{(\rho_w - 1)(\eta - 1)}{\varkappa_{c1}\rho_w - 1}, \kappa_{c\pi} = \frac{(\rho_\pi - 1)[\theta(1 - \eta) + 1]}{\theta(\varkappa_{c1}\rho_\pi - 1)}. \tag{E1.10}$$

And to ensure that the constant terms in (E1.7) equal zero, the coefficient κ_{c0} is calculated from (E1.2) and (E1.4) as, $\kappa_{c0} = \frac{\log \alpha + \varkappa_{c0}}{1-\varkappa_{c1}}$. We then solve for ℓ_c (the unconditional mean of ℓ_{ct}) using Equations (E1.3), E(1.4), (E1.10) and κ_{c0} in the steady state, that is, when $\ell_c = \kappa_{c0} + \kappa_{cw} w + \kappa_{c\pi}\pi$ (where w and π are the steady state values).

Note that the log of the nominal pricing kernel $\omega_{t+1} = \lambda_{t+1} + \pi_t - \pi_{t+1}$ is

$$\omega_{t+1} = B_0 + B_w w_t + (B_\pi + (1-\rho_\pi))\pi_t + b_w \varepsilon_{w,t+1} + b_\pi \varepsilon_{\pi,t+1}. \tag{E1.11}$$

It is also useful to record the log-linearization of the expected nominal SDF around the deterministic steady state. In subsequent analysis, we let $H'_t \equiv \log(H_t) - \log(H)$ be the log deviation of any variable H_t from its steady state value H. Now, in the steady state, $\pi_t = \pi_{t+1}$ and hence $\Omega = \Lambda = \alpha$, so that $\mathbb{E}_t\left[\Omega'_{t,t+1}\right] = \mathbb{E}_t\left[\omega_{t+1} - \alpha\right]$. Therefore, from (E1.11), we have

$$\mathbb{E}_t\left[\Omega'_{t,t+1}\right] = (\theta - 1)\log\alpha + B_w w_t + (B_\pi + (1-\rho_\pi))\pi_t. \tag{E1.12}$$

E2. Loglinearization of Bertrand-Nash equilibrium path

We note that the steady state of the OEP BN equilibrium and the equilibrium follows along the lines of Section C.4 except that W and P are taken as exogenous. Similarly, in perturbations around the steady state, the paths follow as in Section B2 and C2 with the proviso that W_t and P_t are now taken to be exogenous log AR(1) processes (as specified in the text).

Now, log-linearization of the punishment equilibrium dividends (B2.3) gives

$$\begin{aligned}\tilde{D}\tilde{D}'_t \quad &= \quad \tilde{\Upsilon}\left[\frac{W'_t}{\sigma} + \frac{(\sigma-1)}{\sigma}\left(A'_t + P'_t + \psi\tilde{K}'_t\right)\right] - hA\tilde{K}^{\psi}(A'_t + \psi\tilde{K}'_t) - \\ &\quad\quad \delta\tilde{K}\tilde{I}'_t - m\tilde{K}\tilde{K}'_t,\end{aligned} \tag{E2.1}$$

Meanwhile, the derivation of the equilibrium law of motion for capital stock proceeds, in the standard fashion, by developing a second order difference equation in $\tilde{K}'_t$ from log-linearizing the Euler equation (B2.10) through the log-linearization of the capital transition equation $\tilde{I}_t = \tilde{K}_{t+1} - (1-\delta)\tilde{K}_t$, so that

$$\tilde{I}(1 + \tilde{I}'_t) = \tilde{K}[(1 + \tilde{K}'_{t+1}) - (1-\delta)(1 + \tilde{K}'_t)]. \tag{E2.2}$$

But because $\tilde{I} = \delta\tilde{K}$, we have $\tilde{K}/\tilde{I} = 1/\delta$. And since $\tilde{I} = \tilde{K}[1-(1-\delta)]$, (E2.2) yields

$$\tilde{I}'_t = \delta^{-1}\left(\tilde{K}'_{t+1} - (1-\delta)\tilde{K}'_t\right). \tag{E2.3}$$

In particular, (E2.3) implies that $\tilde{I}'_t - \tilde{K}'_t = (\delta)^{-1}(\tilde{K}'_{t+1} - \tilde{K}'_t)$. Then log-linearization of (B2.10) gives, upon noting that $\Omega = \alpha$, canceling out the constant terms due to the steady state Euler equation (E1.12), and taking period t terms under the expectation operator,

$$
\begin{aligned}
0 \;=\;& -\varphi(\tilde{K}'_{t+1} - \tilde{K}'_t) + \mathbb{E}_t \Bigg[\psi \tilde{\Upsilon}^K \Bigg(\alpha \Omega'_{t,t+1} + \frac{W'_{t+1}}{\sigma} + \frac{(\sigma-1)}{\sigma}\left(A'_{t+1} + P'_{t+1}\right) - \\
& y\tilde{K}'_{t+1} \Bigg) - \psi h A \tilde{K}^{(\psi-1)}(\alpha\Omega'_{t,t+1} + A'_{t+1} + (\psi-1)\tilde{K}'_{t+1}) - m + \\
& \alpha\varphi\delta(\tilde{K}'_{t+2} - \tilde{K}'_{t+1})) + (1-\delta)\{\alpha\Omega'_{t,t+1} + \alpha\varphi\delta(\tilde{K}'_{t+2} - \tilde{K}'_{t+1})\}]. \qquad \text{(E2.4)}
\end{aligned}
$$

Using $\tilde{I}'_{t+1} = \delta^{-1}\left(\tilde{K}'_{,t+2} - (1-\delta)\tilde{K}'_{t+1}\right)$ and rearranging terms, the RHS of (E2.4) can be written

$$
\begin{aligned}
0 \;=\;& \tilde{K}'_t \tilde{T}^{;K}_{K0} + \mathbb{E}_t \Big[\Omega'_{t,t+1}\tilde{T}^{;K}_{\Omega} + \tilde{K}'_{t+1}\tilde{T}^{;K}_{K1} + \tilde{K}'_{,t+2}\tilde{T}^{;K}_{2} + \\
& W'_{t+1}\tilde{T}^{;K}_{W1} + P'_{t+1}\tilde{T}^{;K}_{P1} + A'_{t+1}\tilde{T}^{;K}_{A1} \Big], \qquad \text{(E2.5)}
\end{aligned}
$$

where

$$
\begin{aligned}
\tilde{T}^K_{\Omega} &\equiv \alpha\left[\psi(\tilde{\Upsilon}^K - hA\tilde{K}^{(\psi-1)}) - m + (1-\delta)\right], \\
\tilde{T}^K_{K0} &= \varphi, \tilde{T}^K_{K2} = \alpha\varphi\delta(2-\delta), \\
\tilde{T}^K_{K1} &= -\left[\varphi(1+\alpha\delta(2-\delta)) + \alpha\psi(\tilde{\Upsilon}^K y + (\psi-1)hA\tilde{K}^{(\psi-1)}\right], \\
\tilde{T}^K_{W1} &= \alpha\psi\frac{\tilde{\Upsilon}^K}{\sigma}, \tilde{T}^K_{P1} = \alpha\psi\frac{\tilde{\Upsilon}^K(\sigma-1)}{\sigma}, \\
\tilde{T}^K_{A1} &= \alpha\psi\left[\frac{\tilde{\Upsilon}^K(\sigma-1)}{\sigma} - hA\tilde{K}^{(\psi_{-1})}\right]. \qquad \text{(E2.6)}
\end{aligned}
$$

Next, we express $\tilde{k}_{t+1}$ as an affine function of the state variables, that is, of the form

$$\tilde{k}_{t+1} = \tilde{Z}^k_0 + \tilde{Z}^k_w w_t + \tilde{Z}^k_\pi \pi_t + \tilde{Z}^k_a a_t + \tilde{Z}^k_k \tilde{k}_t, \qquad \text{(E2.7)}$$

and determine the coefficients through the Euler condition (E2.5). Note that under the assumed policy in (E2.7),

$$\mathbb{E}_t\left[\tilde{K}'_t\tilde{T}^K_{K0} + \tilde{K}'_{t+1}\tilde{T}^K_{K1} + \tilde{K}'_{t+2}\tilde{T}^K_{K2}\right] =$$

$$
\begin{aligned}
& \tilde{k}_t\left[\tilde{T}^K_{K2}(\tilde{Z}^k_k)^2 + \tilde{T}^K_{K1}\tilde{Z}^k_k + \tilde{T}^K_{K0}\right] + w_t\left[\tilde{Z}^k_w\{\tilde{T}^K_{K2}(\rho_w + \tilde{Z}^k_k) + \tilde{T}^K_{K1}\}\right] + \\
& \pi_t\left[\tilde{Z}^k_\pi\{\tilde{T}^K_{K2}(\rho_\pi + \tilde{Z}^k_k) + \tilde{T}^K_{K1}\}\right] + a_t\left[\tilde{Z}^k_a\{\tilde{T}^K_{K2}(\rho_a + \tilde{Z}^k_k) + \tilde{T}^K_{K1}\}\right] + \\
& \tilde{Z}^k_0\left[\tilde{T}^K_{K2}(1 + \tilde{Z}^k_k) + \tilde{T}^K_{K1}\right] - \tilde{k}[\tilde{T}^K_{K0} + \tilde{T}^K_{K1} + \tilde{T}^K_{K2}]. \qquad \text{(E2.8)}
\end{aligned}
$$

To ensure that terms multiplying $\tilde{k}_t$ equal zero, the following quadratic in $\tilde{Z}_k^k$ must be satisfied

$$\tilde{T}_{K2}^K(\tilde{Z}_k^k)^2 + \tilde{T}_{K0}^K = 0, \tag{E2.9}$$

so that

$$\tilde{Z}_k^k = -\frac{\tilde{T}_{K1}^K}{2T_{K2}^K} \pm \frac{\sqrt{(\tilde{T}_{K1}^K)^2 - 4T_{K2}^K\tilde{T}_{K0}^K}}{2T_{K2}^K}. \tag{E2.10}$$

In the standard way, the smallest real root will be chosen. Next collecting terms for w_t in the Euler (E2.5) and requiring them to be zero, we have

$$w_t\left[\tilde{Z}_w^k\{\tilde{T}_{K2}^K(\rho_w + \tilde{Z}_k^k) + \tilde{T}_{K1}^K\} + \rho_w(\tilde{T}_{W1}^K + \tilde{T}_\Omega B_w)\right] = 0, \tag{E2.11}$$

(where we have used $\mathbb{E}_t[\Omega'_{t,t+1}]$ given in (E1.12)). Hence,

$$\tilde{Z}_w^k = -\frac{\rho_w\tilde{T}_{W1}^K + \tilde{T}_\Omega B_w}{\tilde{T}_{K2}^K(\rho_w + \tilde{Z}_k^k) + \tilde{T}_{K1}^K}. \tag{E2.12}$$

Similarly, we calculate

$$\begin{aligned}\tilde{Z}_\pi^k &= -\frac{\rho_\pi\tilde{T}_{P1}^K + \tilde{T}_\Omega(B_\pi + (1-\rho_\pi))}{\tilde{T}_{K2}^K(\rho_\pi + \tilde{Z}_k^k) + \tilde{T}_{K1}^K}, \\ \tilde{Z}_a^k &= -\frac{\rho_a\tilde{T}_{A1}^K}{\tilde{T}_{K2}^K(\rho_a + \tilde{Z}_k^k) + \tilde{T}_{K1}^K}.\end{aligned} \tag{E2.13}$$

where B_w and B_π are defined in (E1.6) above. Finally, $\tilde{Z}_0^k$ is determined to ensure that the constant terms in (E2.8) sum to zero.

Having determined the coefficients of (E2.7), we similarly express $\tilde{d}_t$ and $\tilde{s}_t$ as affine functions of log of state variables, that is,

$$\tilde{d}_t = \tilde{Z}_0^d + \tilde{Z}_w^d w_t + \tilde{Z}_\pi^d \pi_t + \tilde{Z}_a^d a_t + \tilde{Z}_k^d \tilde{k}_t, \tag{E2.14}$$

$$\tilde{s}_t = \tilde{Z}_0^s + \tilde{Z}_w^s w_t + \tilde{Z}_\pi^s \pi_t + \tilde{Z}_a^s a_t + \tilde{Z}_k^s \tilde{k}_t. \tag{E2.15}$$

Then recognizing that $\tilde{D}'_t = \tilde{d}_t - \tilde{d}K'_t = \tilde{k}_t - \tilde{k}$, and utilizing (E2.1) and (E2.7) yields (up to the constant term)

$$\begin{aligned}
\tilde{Z}^d_w &= \left(\tilde{D}\right)^{-1}\left[\frac{\tilde{\Upsilon}}{\sigma} - \tilde{K}\tilde{Z}^k_w\right], \\
\tilde{Z}^d_\pi &= \left(\tilde{D}\right)^{-1}\left[\tilde{\Upsilon}\left(\frac{\sigma-1}{\sigma}\right) - \tilde{K}\tilde{Z}^k_\pi\right], \\
\tilde{Z}^d_a &= \left(\tilde{D}\right)^{-1}\left[\tilde{\Upsilon}A^{\frac{(\sigma-1)}{\sigma}}\left(\frac{\sigma-1}{\sigma}\right) - hA\tilde{K}^\psi - \tilde{K}\tilde{Z}^k_a\right], \\
\tilde{Z}^d_k &= \left(\tilde{D}\right)^{-1}\tilde{K}\left[(1-\delta) - \tilde{Z}^k_k - m\right].
\end{aligned} \tag{E2.16}$$

Next, we use the log-linearization of the equilibrium asset pricing condition

$$\mathbb{E}_t\left[\Omega_{t,t+1}\left(\frac{\tilde{D}_{t+1}}{\tilde{S}_t} + \frac{\tilde{S}_{t+1}}{\tilde{S}_t}\right)\right] = 1. \tag{C2.35}$$

Loglinear expansion of (C2.35) around the steady state (using $\Omega = \alpha$) gives:

$$\alpha\mathbb{E}_t\left[\left(\frac{\tilde{D}}{\tilde{S}}\right)\left(1 + \Omega'_{t,t+1} + \tilde{D}'_{t+1} - \tilde{S}_t\right) + \left(1 + \Omega'_{t,t+1} + \tilde{S}'_{t+1} - \tilde{S}_t +\right)\right] = 1. \tag{E2.18}$$

However, using the steady state return relationship $\left(1 + \frac{\tilde{D}}{\tilde{S}}\right) = 1/\alpha$, (E2.18) becomes

$$\mathbb{E}_t\left[\Omega'_{t,t+1} + (1-\alpha)\tilde{D}'_{t+1} + \alpha\tilde{S}'_{t+1} - \tilde{S}'_t\right] = 0. \tag{E2.19}$$

Note that

$$\begin{aligned}
\mathbb{E}_t\left[\alpha\tilde{S}'_{t+1} - \tilde{S}'_t\right] &= (\alpha-1)(\tilde{Z}^s_0 - \tilde{s}) + \tilde{Z}^s_w(\alpha\rho_w - 1)w_t + \tilde{Z}^s_\pi(\alpha\rho_\pi - 1)\pi_t + \\
&\quad \tilde{Z}^s_a(\alpha\rho_a - 1)a_t + \tilde{Z}^s_k(\alpha\tilde{k}_{t+1} - \tilde{k}_t),
\end{aligned} \tag{E2.20}$$

$$\begin{aligned}
\mathbb{E}_t\left[(1-\alpha)\tilde{D}'_{t+1}\right] &= (1-\alpha)\left[\tilde{Z}^d_0 + \tilde{Z}^d_w\rho_w w_t + \tilde{Z}^d_\pi\rho_\pi\pi_t + \tilde{Z}^d_a\rho_a a_t + \right. \\
&\quad \left.\tilde{Z}^d_k\tilde{k}_{t+1} - \tilde{d}\right],
\end{aligned} \tag{E2.21}$$

$$\alpha\tilde{k}_{t+1} - \tilde{k}_t = \alpha\left[\tilde{Z}^k_0 + \tilde{Z}^k_w w_t + \tilde{Z}^k_\pi\pi_t + \tilde{Z}^k_a a_t\right] + (\alpha\tilde{Z}^k_k - 1)\tilde{k}_t, \tag{E2.22}$$

$$(1-\alpha)\tilde{k}^*_{t+1} = (1-\alpha)\left[\tilde{Z}^k_0 + \tilde{Z}^k_w w_t + \tilde{Z}^k_\pi\pi_t + \tilde{Z}^k_a a_t + \tilde{Z}^k_k\tilde{k}_t\right]. \tag{E2.23}$$

Thus substituting (E2.22)-(E2.23) in (E2.20)-(E2.21) we get (up to constants),

$$\mathbb{E}_t\left[\Omega'_{t,t+1} + (1-\alpha)\tilde{D}'_{t+1} + \alpha\tilde{S}'_{t+1} - \tilde{S}'_t\right] =$$

$$w_t\left(B_w\rho_w + \tilde{Z}^s_w(\alpha\rho_w - 1) + (1-\alpha)\tilde{Z}^d_w\rho_w + \tilde{Z}^k_w(\alpha\tilde{Z}^s_k + (1-\alpha)\tilde{Z}^d_k)\right) + \pi_t\left(B_\pi\rho_\pi + (1-\rho_\pi) + \tilde{Z}^s_\pi(\alpha\rho_\pi - 1) + (1-\alpha)\tilde{Z}^d_\pi\rho_\pi + \tilde{Z}^k_\pi(\alpha\tilde{Z}^s_k + (1-\alpha)\tilde{Z}^d_k)\right) + a_t\tilde{Z}^s_a(\alpha\rho_a - 1) + (1-\alpha)\tilde{Z}^d_a\rho_a + \tilde{Z}^k_a(\alpha\tilde{Z}^s_k + (1-\alpha)\tilde{Z}^d_k)) + \tilde{k}_t\left(\tilde{Z}^s_k(\alpha\tilde{Z}^k_k - 1) + (1-\alpha)\tilde{Z}^d_k\tilde{Z}^k_k\right). \tag{E2.24}$$

Then to ensure that the coefficient of $\tilde{k}_t$ is zero, we must have

$$\tilde{Z}^s_k = \frac{(1-\alpha)\tilde{Z}^d_k\tilde{Z}^k_k}{1-\alpha\tilde{Z}^k_k}, \tag{E2.25}$$

where $\tilde{Z}^d_k$ and $\tilde{Z}^k_k$ have been computed above. With $\tilde{Z}^s_k$ in hand, we can compute the coefficients of $\tilde{s}_t$ with respect to the other state variables to ensure that the equilibrium asset market condition (E2.19) is satisfied. We have:

$$\begin{aligned}
\tilde{Z}^s_w &= \frac{B_w + (1-\alpha)\tilde{Z}^d_w\rho_w + \tilde{Z}^k_w(\alpha\tilde{Z}^s_k + (1-\alpha)\tilde{Z}^d_k)}{1-\alpha\rho_w}, \\
\tilde{Z}^s_\pi &= \frac{B_\pi + (1-\rho_\pi) + (1-\alpha)\rho_\pi\tilde{Z}^d_\pi + \tilde{Z}^k_\pi(\alpha\tilde{Z}^s_k + (1-\alpha)\tilde{Z}^d_k)}{1-\alpha\rho_\pi}, \\
\tilde{Z}^s_a &= \frac{(1-\alpha)\tilde{Z}^d_a\rho_a + \tilde{Z}^k_a(\alpha\tilde{Z}^s_k + (1-\alpha)\tilde{Z}^d_k)}{1-\alpha\rho_a}.
\end{aligned} \tag{E2.26}$$

E3. Log-linearization of tacit collusion equilibrium path

We will use log-linearization to approximate the equilibrium path using the TCIC (C2.10), the Euler (C2.27) and the asset market equilibrium condition

$$\mathbb{E}_t\left[\Omega_{t,t+1}\left(\frac{D_{t+1}}{S_t} + \frac{S_{t+1}}{S_t}\right)\right] = 1. \tag{E3.1}$$

Note from (C2.13) that, for $i = 0, 1$ (and up to constants)

$$\chi'_{t+i} = -\psi(\sigma-1)u^{\psi}u'_{t+i} + \Theta\left[A'_{t+i} + \psi(K'_{t+i} + (\sigma+1)u'_{t+i}) - W'_{t+i} - (\sigma-1)P'_{t+i}\right], \tag{E3.2}$$

where

$$\Theta \equiv \frac{hN^{\frac{1}{\sigma}}A^{\frac{1}{\sigma}}K^{\frac{\psi}{\sigma}}u^{\frac{\psi(\sigma+1)}{\sigma}}}{M\phi}. \tag{E3.3}$$

We now log-linearize the Euler condition (C2.27). We recall that $I'_{t+1} - K'_{t+1} = (1/\delta)(K'_{t+2} - K'_{t+1})$ and $\tilde{I}'_{t+1} = (1/\delta)(\tilde{K}'_{t+2} - (1-\delta)\tilde{K}'_{t+1})$. Using these facts, log-linearization of (C2.27) gives (up to steady state constants),

$$
\begin{aligned}
0 \quad = \quad & u_t' T_{u0}^K + K_t' T_{K0}^K + A_t' T_{A0}^K + W_t' T_{W0}^K + P_t' T_{P0}^K + \\
& \mathbb{E}_t \left[\Omega_{t,t+1}' T_\Omega^K + K_{t+1}' T_{K1}^K + K_{t+2}' T_{K2}^K + W_{t+1}' T_{W1}^K + P_{t+1}' T_{P1}^K + \right. \\
& \left. A_{t+1}' T_{A1}^K + u_{t+1}' T_{u1}^K + \tilde{K}_{t+2}' T_{\tilde{K}2}^K \right], \qquad \text{(E3.4)}
\end{aligned}
$$

where (using the definitions in (C4.13)-(C4.15)):

$$
\begin{aligned}
T_\Omega^K &= \alpha \left[\left(\Upsilon^K - \psi h A u^\psi K^{(\psi-1)} \right) \left(1 - \chi u^{-\frac{\psi(\sigma-1)}{\sigma}} \right) - \psi \chi [u^{\frac{\psi}{\sigma}} - 1] h A K^{(\psi-1)} + (1-\chi)(1-\delta-m) \right], \\
T_{K0}^K &= \varphi (1-\chi) + \psi \Theta \chi \left(1 + \alpha \left(\left(\Upsilon^K - \psi h A u^\psi K^{(\psi-1)} \right) u^{-\frac{\psi(\sigma-1)}{\sigma}} - \psi (u^{\frac{\psi}{\sigma}} - 1) h A K^{(\psi-1)} - (1-\delta-m) \right) \right), \\
T_{K1}^K &= -\varphi (1-\chi) - \alpha \left[\Upsilon^K y (1 - \chi u^{-\frac{\psi(\sigma-1)}{\sigma}}) + \psi(\psi-1) h A K^{(\psi-1)} (u^{\psi} - 1) + \varphi(\delta(1-\chi) - \right. \\
& \left. (1-\delta)\{(1-\chi) - \delta\chi\psi\Theta\}) \right], \\
T_{K2}^K &= \alpha\varphi[\delta + (1-\chi)(1-\delta)], T_{\tilde{K}2}^K = -\varphi\delta\chi, \\
T_{u0}^K &= (\Theta(\sigma+1) - (\sigma-1)u^\psi)\chi \left[\varphi\delta\psi + \alpha - \left(\Upsilon^K - \psi h A u^\psi K^{(\psi-1)} \right) u^{-\frac{\psi(\sigma-1)}{\sigma}} - \psi (u^{\frac{\psi}{\sigma}} - 1) h A K^{(\psi-1)} + \right. \\
& \left. m - (1-\delta)) \right], \\
T_{u1}^K &= \alpha \left[\psi \Upsilon^K \frac{(\sigma-1)}{\sigma} - \psi h A K^{(\psi-1)} \left(\psi + \chi u^{-\frac{\psi(\sigma-1)}{\sigma}} \left(\frac{\psi(\sigma-1)}{\sigma} u^{\psi} - \psi \right) + \frac{\chi\psi}{\sigma} \right) + \right. \\
& \left. (1-\delta)\varphi\delta\chi(\Theta(\sigma+1) - (\sigma-1)u^\psi) \right], \\
T_{W0}^K &= -\chi\Theta \left[1 + \alpha \left\{ \left(\Upsilon^K - \psi h A u^\psi K^{(\psi-1)} \right) u^{-\frac{\psi(\sigma-1)}{\sigma}} - \psi (u^{\frac{\psi}{\sigma}} - 1) h A K^{(\psi-1)} - (1-\delta-m) \right\} \right], \\
T_{W1}^K &= \alpha \left[(\sigma)^{-1} \Upsilon^K (1 - \chi u^{-\frac{\psi(\sigma-1)}{\sigma}}) - (1-\delta)\varphi\delta\chi\Theta \right], \\
T_{P0}^K &= (\sigma-1) T_{W0}^K, T_{P1}^K = (\sigma-1) T_{W1}^K, T_{A0}^K = -T_{W0}^K, \\
T_{A1}^K &= \alpha \left[\left\{ \left(\frac{\sigma-1}{\sigma} \right) \Upsilon^K - \psi h A u^{\frac{\psi}{\sigma}} K^{(\psi-1)} \right\} (1 - \chi u^{-\frac{\psi(\sigma-1)}{\sigma}}) - \psi\chi[u^{\frac{\psi}{\sigma}} - 1] h A K^{(\psi-1)} + (1-\delta)\varphi\delta\chi\Theta \right].
\end{aligned}
$$

We next log-linearize the TCIC Because, along the equilibrium path at every t, $S_t = \mathbb{E}_t [V_{t+1}(K_{t+1})]$ and $\tilde{S}_t = \mathbb{E}_t \left[\tilde{V}_{t+1}(K_{t+1}) \right]$, the TCIC can be written as

$$
\bar{D}_t - D_t - (S_t - \tilde{S}_t) = 0. \qquad \text{(E3.5)}
$$

Now, $\bar{D}_t - D_t$ equals

$$
\left[M_t \phi N^{-\frac{1}{\sigma}} K_t^{\frac{\psi(\sigma-1)}{\sigma}} A_t^{\frac{(\sigma-1)}{\sigma}} u_t^{-\frac{\psi}{\sigma}} - h A_t K_t^{\psi} \right] (1 - u_t^{\psi}). \qquad \text{(E3.6)}
$$

Hence, log-linearization of (E3.6) around the equilibrium steady state gives (after canceling out constant terms due to the steady state relation of $\bar{D} - D$)

$$\Upsilon\left(\frac{W_t'}{\sigma} + \frac{(\sigma-1)}{\sigma}\left(A_t' + P_t' + \psi K_t'\right) - \frac{\psi}{\sigma}u_t'\right) - hAK^{\psi}(A_t' + \psi K_t') -$$
$$\Upsilon^u\left[\frac{W_t'}{\sigma} + \frac{(\sigma-1)}{\sigma}\left\{A_t' + P_t' + \psi(u_t' + K_t')\right\}\right] + hA\,(uK)^{\psi}\,(A_t' + \psi(K_t' + u_t')). \tag{E3.7}$$

Then, log-linearization of the TCIC (C2.10) by inserting (E3.7) and rearranging terms yields

$$\begin{aligned} u_t' &= (H_0^u)^{-1}\left(SS_t' - \tilde{S}\tilde{S}_t' - (1-u^{\psi})\left[\Upsilon\left(\frac{W_t'}{\sigma} + \frac{(\sigma-1)}{\sigma}P_t'\right) + \right.\right. \\ &\quad \left.\left. A_t'H_a^u + K_t'H_k^u\right]\right), \end{aligned} \tag{E3.8}$$

where

$$\begin{aligned} H_0^u &= \psi\left[hA(uK)^{\psi} - \frac{\Upsilon}{\sigma}\Big(1 + u^{\psi}(\sigma-1)\Big)\right], \\ H_a^u &= \frac{(\sigma-1)}{\sigma}\Upsilon - hAK^{\psi},\ \ H_k^u = \psi H_a^u. \end{aligned} \tag{E3.9}$$

Note that $H_0^u < 0$ since in the steady state $\Upsilon u^{\psi} \geq hA\,(uK)^{\psi}$ (as non-negative profits are necessary for firm survival in steady state). But

$$\Upsilon\left(\frac{1 + u^{\psi}(\sigma-1)}{\sigma}\right) = \Upsilon u^{\psi}\left(\frac{\sigma-1}{\sigma} + \frac{1}{\sigma}\right) + \Upsilon\frac{(1-u^{\psi})}{\sigma} > \Upsilon u^{\psi}. \tag{E3.10}$$

Since (E3.8) holds for any t along the equilibrium path, it follows therefore that

$$\begin{aligned} u_{t+1}' &= (H_0^u)^{-1}\left(SS_{t+1}' - \tilde{S}\tilde{S}_{t+1}' - (1-u^{\psi})\left[\Upsilon\left(\frac{W_{t+1}'}{\sigma} + \frac{(\sigma-1)}{\sigma}P_{t+1}'\right) + \right.\right. \\ &\quad \left.\left. A_{t+1}'H_a^u + K_{t+1}'H_k^u\right]\right). \end{aligned} \tag{E3.11}$$

We can then use (E3.8) and (E3.11) to substitute out u_t' and u_{t+1}' in (E3.4), thereby obtaining an approximation of the optimality condition (E3.4) as an affine function of exogenous state variables $(W_t', W_{t+1}', P_t', P_{t+1}', A_t', A_{t+1}')$ and the costate variables $(K_t',\ K_{t+1}',\ K_{t+2}',\ S_t',\ S_{t+1}',\ \tilde{S}_t',\ \tilde{S}_{t+1}')$. We now use the method of undetermined coefficients to determine the coefficients for the optimality condition by exploiting the asset pricing equilibrium condition:

$$\alpha\mathbb{E}_t\left[\Omega_{t,t+1}' + (1-\alpha)D_{t+1}' + \alpha S_{t+1}' - S_t'\right] = 0. \tag{E3.12}$$

We start by positing equilibrium affine functions for log investment and and log (ex-dividend) equity values, namely,

$$k_{t+1} = Z_0^k + Z_w^k w_t + Z_\pi^k \pi_t + Z_a^k a_t + Z_k^k k_t, \quad \text{(E3.13)}$$

$$s_t = Z_0^s + Z_w^s w_t + Z_\pi^s \pi_t + Z_a^s a_t + Z_k^s k_t. \quad \text{(E3.14)}$$

Using the above two equations, recognizing that $S_t' = s_t - s$, $K_{t+j}' = k_{t+j} - k, j = 0, 1, 2$, etc., using $\Omega_{t,t+1}'$ in (E1.12), and utilizing the fact that $\boldsymbol{\mu}_{t+1}' = \text{diag}(\boldsymbol{\rho})' \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_{t+1}$ yields the log-linearized optimality condition (E3.4) of the firm he form:

$$\begin{aligned} 0 &= \mathbb{E}_t \left[k_t H_{k0}^k + k_{t+1} H_{k1}^k + k_{t+2} H_{k2}^k \right] + w_t H_w^k + H_\pi^k \pi_t + \\ & H_a^k a_t + H_0^k, \end{aligned} \quad \text{(E3.15)}$$

where the coefficients with respect to the state variables are

$$\begin{aligned} H_{k0}^k &= T_{K0}^K + \left(\frac{T_{u0}^K}{H_0^u} \right) \left[S Z_k^s - \tilde{S} \tilde{Z}_k^s - (1 - u^\psi) H_k^u) \right], \\ H_{k1}^k &= T_{K1}^K + \left(\frac{T_{u1}^K}{H_0^u} \right) [S Z_k^s - \tilde{S} \tilde{Z}_k^s - (1 - u^\psi) H_k^u)], \\ H_{k2}^k &= T_{K2}^K, \\ H_w^k &= T_{W0}^K + \rho_w \left(T_{W1}^K + T_\Omega^K B_w \right) + \left(\frac{T_{u0}^K + \rho_w T_{u1}^K}{H_0^u} \right) \times \\ & [S Z_w^s - \tilde{S} \tilde{Z}_w^s - (1 - u^\psi) \frac{\Upsilon}{\sigma}] + (T_{K1}^K + \rho_w T_{K2}^K) Z_w^k, \\ H_\pi^k &= T_{P0}^K + \rho_\pi T_{P1}^K + T_\Omega^K (\rho_\pi (B_\pi - 1) + 1) + \left(\frac{T_{u0}^K + \rho_\pi T_{u1}^K}{H_0^u} \right) \times \\ & \left[S Z_\pi^s - \tilde{S} \tilde{Z}_\pi^s - (1 - u^\psi) \frac{\Upsilon(\sigma - 1)}{\sigma} \right] + (T_{K1}^K + \rho_\pi T_{K2}^K) Z_\pi^k, \\ H_a^k &= T_{A0}^K + \rho_a T_{A1}^K + \left(\frac{T_{u0}^K + \rho_a T_{u1}^K}{H_0^u} \right) \times \\ & [S Z_a^s - \tilde{S} \tilde{Z}_a^s - (1 - u^\psi) H_a^u] + (T_{K1}^K + \rho_a T_{K2}^K) Z_a^k. \end{aligned} \quad \text{(E3.16)}$$

where the terms $T_\cdot^K$ have been defined in (E3.4). However, in (E3.16) H_{k0}^k and H_{k1}^k include (as yet) undetermined stock price loading on capital, Z_k^s, that is the coefficient of k_t in s_t (see (E3.14)).

To determine the coefficients of s_t in general, we utilize the log-linearization of the asset pricing condition in (E3.1), which requires D_t'. Log-linearizing D_t, we have

$$\begin{aligned} D D_t' &= \Upsilon^u \left(\frac{W_t'}{\sigma} + \frac{(\sigma - 1)}{\sigma} \{ P_t' + A_t' + \psi (K_t' + u_t') \} \right) - h A (u K)^\psi \times \\ & (A_t' + \psi (K_t' + u_t')) - \delta K I_t^{'} - m K K_t', \end{aligned} \quad \text{(E3.17)}$$

where $I'_t = (1/\delta)(K'_{t+1} - (1-\delta)K'_t)$. Hence, since $D'_t = d_t - d$, substituting for u'_t from (E3.8) and k_{t+1} from (E3.13), we can write

$$d_t = Z^d_0 + Z^d_w w_t + Z^d_\pi \pi_t + Z^d_a a_t + Z^d_k k_t + Z^d_s s_t + Z^d_{\tilde{s}} \tilde{s}_t, \qquad \text{(E3.18)}$$

where (up to the constant term),

$$\begin{aligned}
Z^d_w &= \frac{1}{\sigma H^u_0 D}\left[H^u_0 \Upsilon^u - \psi(1-u^\psi)\Upsilon\left\{\frac{(\sigma-1)}{\sigma}\Upsilon^u - hA\,(uK)^\psi\right\} - KZ^k_w\right],\\
Z^d_\pi &= \frac{(\sigma-1)}{\sigma H^u_0 D}\left[H^u_0 \Upsilon^u - \psi(1-u^\psi)\Upsilon\left\{\frac{(\sigma-1)}{\sigma}\Upsilon^u - hA\,(uK)^\psi\right\} - KZ^k_\pi\right],\\
Z^d_a &= \frac{\left(\frac{(\sigma-1)}{\sigma}\Upsilon^u - hA\,(uK)^\psi\right)}{H^u_0 D}\left[H^u_0 - \psi(1-u^\psi)H^u_a\right] - KZ^k_a,\\
Z^d_s &= \frac{\psi\left(\frac{(\sigma-1)}{\sigma}\Upsilon^u - hA\,(uK)^\psi\right)S}{H^u_0 D},\\
Z^d_{\tilde{s}} &= -\frac{\psi u^\psi\left(\frac{(\sigma-1)}{\sigma}\Upsilon^u - hA\,(uK)^\psi\right)\tilde{S}}{H^u_0 D},\\
Z^d_k &= \frac{[1-\delta-m-Z^k_k]K}{D} + \frac{\psi\left(\frac{(\sigma-1)}{\sigma}\Upsilon^u - hA\,(uK)^\psi\right)\left[H^u_0 - (1-u^\psi)H^u_k\right]}{H^u_0 D},
\end{aligned} \qquad \text{(E3.19)}$$

and where H^u_0 and H^u_k are defined in (E3.16), while the coefficients $Z^k_\cdot$ of k_{t+1} in (E3.13) will be determined next. We note from (E3.13)-(E3.14) that k_t and s_t do not directly depend on the (log of) the off-equilibrium ex-dividend price $\tilde{s}_t$, while d_t in (E3.18) does so. The reason is apparent from the Euler condition (C2.27) and the equity market clearing condition (E3.1). These conditions do no not directly depend on $\tilde{S}_t$; for example, in the Euler condition off-equilibrium behavior operates through off-equilibrium path investment $\tilde{I}_{t+1}$. However, the optimality condition for capacity utilization u_t in (E3.11) depends directly on $\tilde{S}_t$, which in turn directly affects d_t.

Returning to the asset market equilibrium condition, we can utilize (E3.18)-(E3.19) to substitute for d_{t+1} in the asset pricing condition (E3.12) to derive the equilibrium coefficients for s_t. To this end, note that

$$\alpha k_{t+1} - k_t = \alpha\left[Z^k_0 + Z^k_w w_t + Z^k_\pi \pi_t + Z^k_a a_t\right] + (\alpha Z^k_k - 1)k_t, \qquad \text{(E3.20)}$$

$$(1-\alpha)k_{t+1} = (1-\alpha)\left[Z^k_0 + Z^k_w w_t + Z^k_\pi \pi_t + Z^k_a a_t + Z^k_k k_t\right]. \qquad \text{(E3.21)}$$

Hence, collecting terms in (the LHS) of (E3.12), the asset pricing condition requires

$$\begin{aligned}
0 \;=\; & w_t \Big[(1-\alpha)\rho_w(Z_w^d + Z_{\tilde{s}}^d \tilde{Z}_w^s) + Z_w^s\{(\alpha\rho_w - 1) + (1-\alpha)\rho_w Z_s^d\} + \\
& Z_w^k\{(1-\alpha)Z_k^d + \alpha Z_k^s\} + B_w\Big] + \pi_t \Big[(1-\alpha)\rho_\pi(Z_\pi^d + Z_{\tilde{s}}^d \tilde{Z}_\pi^s) + Z_\pi^s\{(\alpha\rho_\pi - 1) + \\
& (1-\alpha)\rho_\pi Z_s^d\} + Z_\pi^k\{(1-\alpha)Z_k^d + \alpha Z_k^s\} + B_\pi + (1-\rho_\pi)\Big] + \\
& a_t \Big[(1-\alpha)\rho_a(Z_a^d + Z_{\tilde{s}}^d \tilde{Z}_a^s) + Z_a^s\{(\alpha\rho_a - 1) + (1-\alpha)\rho_a Z_s^d\} + \\
& Z_a^k\{(1-\alpha)Z_k^d + \alpha Z_k^s\}\Big] + k_t \Big[(1-\alpha)(Z_k^d + Z_{\tilde{s}}^d \tilde{Z}_k^s) + Z_k^s(\alpha Z_k^k - 1)\Big] + Z_0^s. \qquad \text{(E3.22)}
\end{aligned}$$

It is clear that to ensure that the coefficient of k_t is zero for all values of the state, we must have

$$Z_k^s = \frac{(1-\alpha)(Z_k^d + Z_{\tilde{s}}^d \tilde{Z}_k^s)}{1-\alpha Z_k^k}. \qquad \text{(E3.23)}$$

But we notice from (E3.19) that

$$\begin{aligned}
Z_k^d &= H_k^d - K Z_k^k, \\
H_k^d &\equiv \frac{(1-\delta-m)K}{D} + \frac{\psi(\Upsilon - hA\,(uK)^\psi)\left[H_0^u - (1-u^\psi)H_k^u\right]}{H_0^u D}. \qquad \text{(E3.24)}
\end{aligned}$$

Hence, substituting (E3.24) in (E3.23) and then substituting the resultant expression for Z_k^s in (E3.16), we get

$$\begin{aligned}
H_{k0}^K &= \left(\frac{1}{1-\alpha Z_k^k}\right)[\Xi_{00}^K + Z_k^k \Xi_{01}^K], \\
H_{k1}^K &= \left(\frac{1}{1-\alpha Z_k^k}\right)[\Xi_{10}^K + Z_k^k \Xi_{11}^K], \qquad \text{(E3.25)}
\end{aligned}$$

where

$$\begin{aligned}
\Xi_{00}^K &= T_{K0}^K + \left(\frac{T_{u0}^K}{H_0^u}\right)[S(1-\alpha)H_k^d + \tilde{S}\tilde{Z}_k^s((1-\alpha)Z_{\tilde{s}}^d - 1) - (1-u^\psi)H_k^u)], \\
\Xi_{01}^K &= -\alpha T_{K0}^K - \left(\frac{T_{u0}^K}{H_0^u}\right)[SK(1-\alpha) - \alpha(\tilde{S}\tilde{Z}_k^s + (1-u^\psi)H_k^u)], \\
\Xi_{10}^K &= T_{K1}^K + \left(\frac{T_{u1}^K}{H_0^u}\right)[S(1-\alpha)H_k^d + \tilde{S}\tilde{Z}_k^s((1-\alpha)Z_{\tilde{s}}^d - 1) - (1-u^\psi)H_k^u)], \\
\Xi_{11}^K &= -\alpha T_{K1}^K - \left(\frac{T_{u1}^K}{H_0^u}\right)[SK(1-\alpha) - \alpha(\tilde{S}\tilde{Z}_k^s + (1-u^\psi)H_k^u)]. \qquad \text{(E3.26)}
\end{aligned}$$

Returning to (E3.15)-(E3.16), to ensure that the terms involving the log of the capital stock are collectively zero, Z_k^k is derived as the smallest real root of the cubic equation:

$$\hat{H}_{k0}^K + \hat{H}_{k1}^K Z_k^k + \hat{H}_{k2}^K (Z_k^k)^2 + \hat{H}_{k3}^K (Z_k^k)^3 = 0, \tag{E3.27}$$

where

$$\begin{aligned} \hat{H}_{k0}^K &= \Xi_{00}^K, \\ \hat{H}_{k1}^K &= \Xi_{01}^K + \Xi_{10}^K, \\ \hat{H}_{k2}^K &= H_{k2}^K + \alpha \Xi_{11}^K, \\ \hat{H}_{k3}^K &= -\alpha H_{k2}^K. \end{aligned} \tag{E3.28}$$

Note that the determination of Z_k^k (in (E3.27)-(E3.28)) also yields Z_k^d (from (E3.19)) and Z_k^s (from (E3.23)). In a similar fashion, we can compute the loadings of k_{t+1} of s_t on the other state variables (w_t, π_t, a_t) by utilizing the Euler condition (E3.15) and the asset market condition (E3.22), which then yield the corresponding loadings of d_t (from (E3.19)). More explicitly, because we require the coefficients of the state variables in the asset market condition (E3.22) to each be zero, we can solve for the loadings of s_t as

$$\begin{aligned} Z_w^s &= \frac{(1-\alpha)\rho_w (Z_w^d + Z_{\tilde{s}}^d \tilde{Z}_w^s) + Z_w^k \{(1-\alpha) Z_k^d + \alpha Z_k^s\} + B_w}{(1-\alpha\rho_w) + (\alpha - 1)\rho_w Z_s^d}, \\ Z_\pi^s &= \frac{(1-\alpha)\rho_\pi (Z_\pi^d + H_{\tilde{s}}^d \tilde{Z}_\pi^s) + Z_\pi^k \{(1-\alpha) Z_k^d + \alpha Z_k^s\} + B_\pi + (1-\rho_\pi)}{(1-\alpha\rho_\pi) + (\alpha - 1)\rho_\pi Z_s^d}, \\ Z_a^s &= \frac{(1-\alpha)\rho_a (Z_a^d + Z_{\tilde{s}}^d \tilde{Z}_a^s) + Z_a^k \{(1-\alpha) Z_k^d + \alpha Z_k^s\}}{(1-\alpha\rho_a) + (\alpha - 1)\rho_a Z_s^d}. \end{aligned} \tag{E3.29}$$

Finally, we utilize the foregoing derivations to compute nominal equity returns (for the representative firm in any industry) by using the Campbell and Shiller (1988) approximation

$$r_{t+1} = \varkappa_{d0} + \varkappa_{d1} \ell_{d,t+1} - \ell_{dt} + d_{t+1} - d_t, \tag{E3.30}$$

where ℓ_{dt} is the log price-dividend ratio $\log(S_t/D_t) = s_t - d_t$. And $\varkappa_{d0}$ and $\varkappa_{d1}$ are given by

$$\varkappa_{d0} = \log[1 + \exp(\ell_d)] - \varkappa_{d1} \ell_d, \varkappa_{d1} = \frac{\exp(\ell_d)}{1 + \exp(\ell_d)}, \tag{E3.32}$$

where ℓ_d is the unconditional mean of ℓ_{dt}. From the foregoing, it follows that

$$\ell_{dt} = \kappa_{d0} + \kappa_{dw} w_t + \kappa_{d\pi} \pi_t + \kappa_{da} a_t + \kappa_{dk} k_t + \kappa_{d\tilde{s}} \tilde{s}_t, \tag{E3.33}$$

where (using (E3.14) and (E3.18)):

$$\begin{aligned}\kappa_{d0} &= \frac{Z_0^s - Z_0^d}{1+Z_d^s}; \kappa_{dw} = \frac{Z_w^s - Z_w^d}{1+Z_d^s}; \kappa_{d\pi} = \frac{Z_\pi^s - Z_\pi^d}{1+Z_d^s}; \kappa_{da} = \frac{Z_a^s - Z_a^d}{1+Z_d^s};\\ \kappa_{dk} &= \frac{Z_k^s - Z_k^d}{1+Z_d^s}; \kappa_{d\tilde{s}} = \frac{-Z_{\tilde{s}}^d}{1+Z_d^s}.\end{aligned} \tag{E3.34}$$

In our numerical simulations of conditional returns r_{t+1}, we take ℓ_d to be the mean of the simulated $(s_t - d_t)$, using the equilibrium rules derived above. ∎

F. Depreciation costs of capacity utilization

We now consider the case where $\delta(u) = \frac{1}{\xi}u^\xi, \xi > 1$. The steady state capital adjustment cost function is $\Psi(I,K) = I + 0.5\varphi\left(\frac{I}{K} - \delta(u)\right)^2 K$. It is convenient to define

$$\begin{aligned}\tilde{\Upsilon} &= M\phi N^{-\frac{1}{\sigma}} A^{\frac{(\sigma-1)}{\sigma}} (\tilde{u}\tilde{K})^{\frac{\psi(\sigma-1)}{\sigma}},\\ \tilde{\Upsilon}^K &= \left[M\phi N^{-\frac{\sigma+1}{\sigma}} A^{\frac{\sigma-1}{\sigma}} \tilde{u}^{\frac{\psi(\sigma-1)}{\sigma}} \tilde{K}^{-y}\left(\frac{N\sigma-1}{\sigma}\right)\right],\\ \tilde{\Upsilon}^u &= \left[M\phi A^{\frac{\sigma-1}{\sigma}} N^{-\frac{\sigma+1}{\sigma}} \tilde{u}^{-y} \tilde{K}^{\frac{\psi(\sigma-1)}{\sigma}}\left(\frac{N\sigma-1}{\sigma}\right)\right],\end{aligned} \tag{F1.1}$$

$y \equiv \frac{(\sigma(1-\psi)+\psi)}{\sigma}$. And when used without time subscripts—that is, $\tilde{\Upsilon}$, $\tilde{\Upsilon}^K$ and $\tilde{\Upsilon}^u$—the notation will refer to the steady state values of these variables. The firm's optimization problem, conditional on capital stock $\tilde{K}$ and other firms' (symmetric) strategies $(\tilde{u}_{-n}, \tilde{K}_{-n})$, can be represented as

$$V(\tilde{K}, \tilde{K}_{-n}) = \max_{u\in[0,1],I} D(u, I \,|\, \tilde{u}_{-n}, \tilde{K}_{-n}, \tilde{K}) + \alpha V(\tilde{K}((1-\delta(u)) + I), \tag{F1.2}$$

where

$$\begin{aligned}\Pi(u, \tilde{u}_{-n}, \tilde{K}_{-n}, \tilde{K}) &= \left(M\phi A^{-\frac{1}{\sigma}}[(N-1)(\tilde{u}\tilde{K})^\psi + (u\tilde{K})^\psi]^{-\frac{1}{\sigma}} - h\right) A(u\tilde{K})^\psi, &(\text{F1.3})\\ D(u, I, \tilde{u}_{-n}, \tilde{K}_{-n}, \tilde{K}) &= \Pi(u \,|\, \tilde{u}_{-n}, \tilde{K}_{-n}, \tilde{K}) - \Psi(I, \tilde{K}) - m\tilde{K}. &(\text{F1.4})\end{aligned}$$

Then, $(\tilde{u}, \tilde{K})$ is the steady state equilibrium for all if the solution to (F1.2)-(F1.4) implies $\tilde{I} = \tilde{\delta}\tilde{K}$, where $\tilde{\delta} = \frac{1}{\xi}(\tilde{u})^\xi$. Thus, $(\tilde{u}, \tilde{K},)$ are, respectively, characterized by the optimality conditions

$$\begin{aligned}0 &= \frac{\partial \Pi(\tilde{u}, \tilde{K})}{\partial u} - \alpha(\tilde{u})^{(\xi-1)}\frac{\partial \tilde{V}(\tilde{K})}{\partial K}, &(\text{F1.5})\\ 0 &= -1 + \alpha\left[\frac{\partial \Pi(\tilde{u}, \tilde{K})}{\partial K} - m + (1-\tilde{\delta})\right] &(\text{F1.6})\end{aligned}$$

Now, from (F1.3)

$$\begin{aligned}\frac{\partial \Pi(\tilde{u},\tilde{K})}{\partial u} &= \frac{\partial \Pi(u,K\,|\,(\tilde{u}_{-n},\tilde{K}_{-n},\tilde{K}))}{\partial u}\Big|_{u=\tilde{u},K=\tilde{K}} \\ &= \psi\left[\frac{(\sigma-1)}{\sigma}\tilde{\Upsilon}^u - hA\tilde{u}^{(\psi-1)}\tilde{K}^{\psi}\right]. \end{aligned} \tag{F1.7}$$

Similarly,

$$\frac{\partial \Pi(\tilde{u},\tilde{K})}{\partial K} = \psi\left[\frac{(\sigma-1)}{\sigma}\tilde{\Upsilon}^K - hA(\tilde{u})^{\psi}(\tilde{K})^{(\psi-1)}\right]. \tag{F1.8}$$

And dividends in the punishment steady state are:

$$\tilde{D} = \Pi(\tilde{u},\tilde{K}) - \Psi(\tilde{\delta}\tilde{K},\tilde{K}) - m. \tag{F1.9}$$

(F1.7)-(F1.9) determine the steady state cum- and ex-dividend equity values as $\tilde{V} = \frac{\tilde{D}}{1-\alpha}$, $\tilde{S} = \frac{\alpha\tilde{D}}{1-\alpha}$.

In the standard fashion, we can use the implicit function theorem on the optimality conditions $\frac{\partial \tilde{V}(\tilde{K})}{\partial u} = 0$ (F1.5) and $\frac{\partial \tilde{V}(\tilde{K})}{\partial I} = 0$ (F1.6) to derive comparative statics with respect to the depreciation cost parameter ξ. Note that (F1.6) implies $1 = \alpha\frac{\partial \tilde{V}(\tilde{K})}{\partial K}$. Substituting this in (F1.5), we compute

$$\frac{\partial^2 \tilde{V}(\tilde{K})}{\partial u \partial \xi} = -\log(u)(\tilde{u})^{(\xi-1)}\tilde{K} > 0, \tag{F1.10}$$

since $\log(u) < 0$ (as $0 < u \leq 1$). And from (F1.6) we have

$$\frac{\partial^2 \tilde{V}(\tilde{K})}{\partial I \partial \xi} = -\alpha\frac{\partial \tilde{\delta}}{\partial \xi} > 0. \tag{F1.11}$$

Next, from Young's Theorem, we have $\frac{\partial^2 \tilde{V}(\tilde{K})}{\partial u \partial K} = \frac{\partial^2 \tilde{V}(\tilde{K})}{\partial K \partial u}$ and hence from (F1.7) and (F1.5)

$$\frac{\partial^2 \tilde{V}(\tilde{K})}{\partial K \partial u} = \frac{\psi^2}{\tilde{K}}\left[\left(\frac{\sigma-1}{\sigma}\right)^2\tilde{\Upsilon}^u - hA\tilde{u}^{(\psi-1)}\tilde{K}^{\psi}\right] - \alpha\tilde{u}^{(\xi-1)}\frac{\partial^2 \tilde{V}(\tilde{K})}{(\partial K)^2}. \tag{F1.12}$$

Then using standard comparative statics techniques for multivariate optimization we get

$$\frac{\partial \tilde{K}}{\partial \xi} = \frac{-\frac{\partial^2 \tilde{V}(\tilde{K})}{(\partial K)^2}\frac{\partial^2 \tilde{V}(\tilde{K})}{\partial I \partial \xi} + \frac{\partial^2 \tilde{V}(\tilde{K})}{\partial K \partial u}\frac{\partial^2 \tilde{V}(\tilde{K})}{\partial u \partial \xi}}{\frac{\partial^2 \tilde{V}(\tilde{K})}{(\partial K)^2}\frac{\partial^2 \tilde{V}(\tilde{K})}{(\partial u)^2} - \left(\frac{\partial^2 \tilde{V}(\tilde{K})}{\partial K \partial u}\right)^2}. \tag{F1.13}$$

Now the local second order conditions for an optimum, $\frac{\partial^2 \tilde{V}(\tilde{K})}{(\partial K)^2} < 0$ and the denominator of (F1.13) is positive. Hence, from (F1.10) and (F1.11), $\frac{\partial \tilde{K}}{\partial \xi} > 0$ if $\frac{\partial^2 \tilde{V}(\tilde{K})}{\partial K \partial u} \geq 0$. Returning to (F1.12), the second term is positive by the local second order conditions. And from (F1.5), $\frac{\partial \Pi(\tilde{u},\tilde{K})}{\partial u} > 0$, that is, (F1.7) is positive. Hence a sufficient condition for $\frac{\partial^2 \tilde{V}(\tilde{K})}{\partial K \partial u} \geq 0$ is that

$$\begin{aligned} h &\leq \left(\frac{\sigma-1}{\sigma}\right)^2 M\phi N^{-\frac{\sigma+1}{\sigma}} A^{\frac{-1}{\sigma}} (\tilde{u}\tilde{K})^{-\frac{\psi}{\sigma}} \left(\frac{N\sigma-1}{\sigma}\right) \\ &= \left(\frac{\sigma-1}{\sigma}\right)^2 N \left(\frac{N\sigma-1}{\sigma}\right) \tilde{p}. \end{aligned} \tag{F1.14}$$

∎

References

Campbell, J., and R. Shiller. 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1:195-228.

Epstein, L., and S. Zin. 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica* 57:937-969.

Kreps, D., and J. Scheinkman. 1983. Quantity precommitment and Bertrand competition yields Cournot outcomes. *Rand Journal of Economics* 14:326–377.

Table A.1. Calibration for quantitative analysis in Section 5.1

This table displays the parameterization used for the numerical computations of the loglinear approximation of the industry equilibrium presented in Section 5.1. ρ_w, ρ_π are the estimated autocorrelation coefficients of the first order autoregressive processes of $w_t = \log W_t$ and $\pi_t = \log\ P_t$, that is, $w_t = \rho_w w_{t-1} + \varepsilon_{wt}$ and $\pi_t = \rho_\pi \pi_{t-1} + \varepsilon_{\pi t}$. W_t and P_t are described in Section 4.1. Σ_w, Σ_π and $\Sigma_{w\pi}$ are the volatilities and covariance of the estimated income and price index shocks ε_{wt} and $\varepsilon_{\pi t}$. The autocorrelation coefficients of sectoral log productivity shock (ρ_a), as well as the covariances of industry productivity shocks with w_t and π_t ($\Sigma_{aw}, \Sigma_{a\pi}, \Sigma_{xw}, \Sigma_{x\pi}$) are estimated from the first order autoregressive processes of log productivity in the NBER-CES (annual) data for U.S. manufacturing consumer goods sector during 1958-2016.

Consumption and Production	
Annual discount factor (α)	0.99
Intertemporal elasticity of substitution (η^{-1})	1.9
Risk aversion (γ)	5
Product elasticity of substitution (σ)	2.5
Sectoral weight (ϕ)	0.05
Output elasticity of capital (ψ)	0.78
Capital adjustment cost (φ)	5
Annual depreciation rate (δ)	5.5%
Variance-Covariance Matrix Σ (Annual)	
Volatility of ε_{wt} (Σ_w)	2.5%
Volatility of $\varepsilon_{\pi t}$ (Σ_π)	2.85%
Volatility of ε_{at} (Σ_a)	0.5%
$Cov(\varepsilon_{w_t}, \varepsilon_{\pi t})$ ($\Sigma_{w\pi}$)	0.001
$Cov(\varepsilon_{w_t}, \varepsilon_{a_t})$ (Σ_{aw})	0.0002
$Cov(\varepsilon_{\pi_t}, \varepsilon_{a_t})$ ($\Sigma_{a\pi}$)	0.0005
Autocorrelation Coefficients (Annual)	
Coefficient of w_t (ρ_w)	0.95
Coefficient of π_t (, ρ_π)	0.965
Coefficient of a_t (ρ_a)	0.88

Table A.2. Time trends in characteristics

This table provides summary statistics—mean, standard deviation (S.D.), and median—of salient variables for the first half of the sample period (1989-2004) and the full sample period (1989-2019). The t-statistics of differences in means of (first half minus full sample) and the P-value of differences in medians (first half minus full sample) are also provided.

	1989-2004			1989-2019			Statistical Tests			
	Mean	S.D.	Median	Mean	S.D.	Median	Diff. in Means	t-stat	Diff. in Medians	P-value
HHI	0.299	0.284	0.183	0.303	0.282	0.189	-0.004	-4.25	-0.006	0.000
Fixed Cost	0.328	0.207	0.285	0.334	0.220	0.286	-0.006	-8.22	-0.001	0.018
Markup	1.334	0.877	1.263	1.396	1.053	1.286	-0.063	-18.75	-0.023	0.000
Markup Volatility	0.187	0.231	0.107	0.185	0.252	0.097	0.001	1.81	0.010	0.000
Capital Investment	13.526	40.383	1.690	26.375	87.057	2.768	-12.849	-51.49	-1.078	0.000
Capital Investment Volatility	0.379	0.229	0.343	0.385	0.229	0.348	-0.007	-8.29	-0.005	0.000
Stock Return	1.827	45.311	0.000	1.474	35.078	0.220	0.354	5.24	-0.220	0.030
Stock Return Volatility	0.627	0.354	0.548	0.557	0.321	0.476	0.070	120.30	0.072	0.000
Net Cash Flow Volatility	18.182	31.457	5.926	36.837	76.366	9.878	-18.655	-77.74	-3.952	0.000
Scaled Net Cash Flow Volatility	0.055	0.031	0.048	0.048	0.029	0.040	0.008	67.50	0.008	0.000