The Oligopoly Lucas Tree

Winston Wei Dou Yan Ji Wei Wu*

May 31, 2021

Abstract

This paper proposes a novel quantitative framework with endogenous strategic competition in heterogeneous concentrated industries. Oligopolies compete strategically for profit margins in repeated games, trading off the benefits of future cooperation against those of reaping higher short-run profits by undercutting their rivals. Cross-industry dispersions in market leadership persistence and cash flow exposures to expected growth simultaneously determine the relationships among profitability, book-to-market ratios, and systematic risk exposures, thereby quantitatively rationalizing the gross profitability and value premium across industries and, importantly, their interactions — controlling for the book-to-market ratio (gross profitability) makes the gross profitability (value) premium more pronounced.

Keywords: Endogenous competition, Strategic rivalry, Value premium, Gross profitability premium, External habit formation. (**JEL**: G12, L13, O33, C73)

*Dou: University of Pennsylvania (wdou@wharton.upenn.edu). Ji: HKUST (jiy@ust.hk). Wu: Texas A&M (wwu@mays.tamu.edu). We are especially grateful to Hengjie Ai, Frederico Belo, Bob Dittmar, Itamar Drechsler, Itay Goldstein, Joao Gomes, Bryan Kelly, Lukas Schmid, and Jessica Wachter for detailed discussions and comments. We also thank Jules van Binsbergen, M. Cecilia Bustamante, Hui Chen, Zhi Da, Lorenzo Garlappi, Xavier Giroud, Vincent Glode, Lars Peter Hansen, David Hirshleifer, Harrison Hong, Urban Jermann, Ken Judd, Ron Kaniel, Don Keim, Richard Kihlstrom, Leonid Kogan, Jun Li, Xiaoji Lin, Andrey Malenko, Gustavo Manso, Robert Novy-Marx, Alan Moreira, Thien Nguyen, Christian Opp, Marcus Opp, Jonathan Parker, Ľuboš Pástor, Thomas Philippon, Raghu Rajan, Uday Rajan, Krishna Ramaswamy, Michael Roberts, Nick Roussanov, Larry Schmidt, Stephanie Schmitt-Grohé, Enrique Schroth, Rob Stambaugh, Luke Taylor, David Thesmar, Robert Townsend, Harald Uhlig, Gianluca Violante, Pengfei Wang, Neng Wang, Wei Xiong, Amir Yaron, and Lu Zhang, and participants at University of Rochester (Simons), Federal Reserve Bank at Dallas, University of Texas at Dallas, City University of Hong Kong, Peking University (Guanghua and PHBS), Wharton, CICM, COAP Conference, European Finance Association, Midwest Finance Association, Hong Kong University of Science and Technology, HKUST Finance Symposium, HKUST-JINAN Workshop, Mack Institute Workshop, MIT Junior Faculty Conference, Mays Innovation Research Center Workshop, PNC Kentucky Finance Conference, Northeastern Finance Conference, 6th SAFE Asset Pricing Workshop, SFS Cavalcade North America, Stanford SITE, and Young Scholars Finance Consortium (YSFC). This paper is partly based on the manuscript previously circulated under the title "Competition, Profitability, and Risk Premia." We are also grateful for the comments and insights from Britt Harris, the President, CEO, and CIO of the University of Texas/Texas A&M Investment Company (UTIMCO).

1 Introduction

This paper proposes a quantitative equilibrium model that extends the standard Lucas-tree asset pricing framework to reserve an explicit role for endogenous strategic competition in heterogeneous concentrated industries.[1] By calibrating the model to the data without using information from stock returns, we show that our model can quantitatively reconcile the gross profitability and value premium across industries, especially their intriguing interactions. In particular, after controlling for the book-to-market ratio and gross profitability, the gross profitability and value premium become more pronounced, respectively. Jointly addressing these two "anomalies" and their interactions is challenging because profitable firms are associated with low book-to-market ratios and are therefore likely to generate a counterfactual "growth premium" (Novy-Marx, 2013). Moreover, the interaction patterns of the two anomalies cannot be easily rationalized by single-factor models, especially in the presence of a strong correlation between gross profitability and the book-to-market ratio across industries.

Many capital market anomalies are firm-level phenomena that are likely to vanish at the industry level. However, the gross profitability and value premium, as well as their interactions, survive across industries. In particular, we show that, in the data, the patterns survive across industries at the level of the four-digit Standard Industrial Classification (SIC4) codes, with statistical significance and economic magnitude comparable to the same patterns across firms within industries and those across all firms. Rationalizing industry-level anomalies adds another layer of difficulty — in general, financial economists find it difficult to connect industry-level risk exposures to the fundamental characteristics of an industry (e.g., Fama and French, 1997; Dittmar and Lundblad, 2017), and the theory of a firm for explaining stock returns cannot be directly extended to that of an industry (e.g., Dou, Ji and Wu, 2021). To connect stock returns and fundamental characteristics across industries,

[1]The closest quantitative general-equilibrium models with heterogeneous firms in an endowment economy include those proposed by Menzly, Santos and Veronesi (2004) and Santos and Veronesi (2006, 2010), who adopt a similar top-down modeling approach. Our paper differs from these studies by explicitly highlighting the crucial role of endogenous strategic competition in asset pricing.

several key industrial organizational features must be explicitly considered: (i) industries are highly concentrated; (ii) the market leadership of major players is persistent; and (iii) market leaders compete strategically for market share within an industry.[2] Despite these features, the asset pricing literature has paid little attention to the effect of endogenous strategic competition. This paper fills this gap by taking a first step toward the development of a full-fledged quantitative asset pricing framework with endogenous strategic competition.

Our explanation of the joint patterns of industry-level stock returns emphasizes the endogeneity of strategic competition and its interplay with two dimensions of cross-industry heterogeneity. In the proposed economic mechanism of endogenous competition, market leaders in a given industry compete strategically through tacit collusion or cooperation sustained by punishment for deviation in repeated games. They optimally set profit margins by trading off the long-run benefits of tacitly colluding with their rivals against the short-run benefits of reaping higher profits by departing from the collusive agreement and undercutting their rivals. The incentive to collude (i.e., collusion capacity) is endogenously driven by fluctuations in the discount rate and expected growth. Both a rise in the discount rate and a decline in expected growth reduce the present value of future cooperation, causing firms to compete more fiercely for short-run profits by undercutting each other. Moreover, this endogenous competition mechanism interacts with primitive industry characteristics in important ways. In the model, cross-industry heterogeneity is introduced through cross-sectional differences in two primitive industry characteristics, the market leadership turnover rate and the cash flow loading on expected growth. Cross-sectional differences in the market leadership turnover rate reflect differences in the average life span of market leaders (i.e., the effective patience of market leaders) in an industry, while cross-sectional differences in the cash flow loading on expected growth reflect differences in the riskiness of firms' growth options in an industry. Intuitively, the dispersion of the market leadership turnover rate, together with discount rate shocks, generates the gross profitability premium, as industries

[2]See, for example, Grullon, Larkin and Michaely (2018), Gutiérrez, Jones and Philippon (2019), Autor et al. (2020), Loecker, Eeckhout and Unger (2020), Corhay, Kung and Schmid (2020*b*), and Dou, Ji and Wu (2021), for evidence related to industry concentration; see, for example, Sutton (2007) and Bronnenberg, Dhar and Dubé (2009) for evidence related to market leadership persistence; and see Chen et al. (2020) and the references therein for evidence related to strategic competition such as tacit collusion.

with lower market leadership turnover rates are effectively more patient; thus, they have stronger endogenous competition effects — their market leaders have the capacity to collude on higher profit margins while being more negatively exposed to discount rate shocks. This is the endogenous competition channel for the industry-level gross profitability premium proposed by Dou, Ji and Wu (2021). Meanwhile, the dispersion of the cash flow loading on expected growth, together with expected growth shocks, generates the value premium, as industries with higher cash flow loadings on expected growth have riskier growth options and thus have higher book-to-market ratios while being more positively exposed to expected growth shocks. This is the cash flow duration channel for the value premium similar in spirit to that proposed by Campbell and Vuolteenaho (2004), Lettau and Wachter (2007, 2011), Da (2009), Santos and Veronesi (2010), and Croce, Lettau and Ludvigson (2014), among others.[3]

At first glance, it may not be very surprising that, by combining two cross sections and two systematic shocks, our model is able to quantitatively explain two anomalies simultaneously in a unified framework. However, it should be noted that the goal of our model goes far beyond merely generating the gross profitability and value premium separately in their own cross sections. Importantly, the main contribution of our model is to advance our understanding of the important and nontrivial interactions between the two cross sections. In doing so, this paper differs considerably from Dou, Ji and Wu (2021) and contributes to the literature in three respects. First, our model shows how gross profitability and book-to-market ratios are jointly determined by two primitive industry characteristics, the market leadership turnover rate and the cash flow loading on expected growth. Second, our model shows how industry-level stock return exposures to these two systematic shocks are jointly determined by the two primitive industry characteristics. Third, our model establishes the cross-industry interdependence between gross profitability, book-to-market ratios, and risk exposures, thereby rationalizing the interactions between the gross profitability and value premium across industries.

The seemingly complex interdependence between two industry characteristics, two

[3]We discuss in detail the connection between our mechanism for generating the value premium and the mechanisms in the literature in Appendix A.

financial ratios, risk exposures to two systematic shocks, and stock returns can be explained in a fairly transparent way because we show that a "nearly separating property" holds in our model, which is also verified in the data. More precisely, we show that cross-industry dispersions in gross profitability and exposure to discount rate shocks, as well as their association, are mainly determined by the dispersion of the market leadership turnover rate. Meanwhile, cross-industry dispersions in the book-to-market ratio and exposure to expected growth shocks, as well as their association, are mainly determined by the dispersion of the cash flow loading on expected growth. This "nearly separating property" is not obvious ex ante. As a main contribution, our model quantitatively confirms this property and uses it to guide our empirical tests of the main theoretical results in the data. The "nearly separating property" of our model ensures that the cross-industry correlation between gross profitability and book-to-market ratios, as well as the interaction between the gross profitability and value premium, ultimately boil down to the correlation parameter of the two primitive industry characteristics. By calibrating the correlation parameter to match its empirical counterpart measured directly in the data, we show that our model can quantitatively match the data in many dimensions, including the complex interactions between the two cross sections of industries.

Specifically, in the data, we find that the returns of more profitable industries are significantly more exposed to discount rate shocks but not to expected growth shocks, whereas the returns of industries with higher book-to-market ratios are more exposed to expected growth shocks but not to discount rate shocks (see Online Appendix 2.4). These empirical findings suggest that the industry-level gross profitability premium is primarily explained by heterogeneous exposure to discount rate shocks, whereas the industry-level value premium is mainly explained by heterogeneous exposure to expected growth shocks, consistent with the "nearly separating property" implied by our model.

We also conduct counterfactual analyses to shed light on the importance of each ingredient in our model. The model implies that the time-varying discount rate plays a major role in generating the gross profitability premium. By contrast, fluctuations in expected growth are necessary and important to generate the value premium. We quantitatively isolate the contribution of the two industry characteristics. The dispersion of gross profitability across industries is mainly attributed to the difference in the market leadership turnover rate. The model cannot generate a gross profitability return spread if the turnover rate of market leaders is the same across industries. Meanwhile, the dispersion of book-to-market ratios across industries is mainly attributed to the difference in the cash flow loading on expected growth. Without dispersion in the cash flow loading on expected growth across industries, the model would generate a growth premium rather than a value premium. Finally, we show that the correlation between the two primitive characteristics across industries plays a key role in shaping the interactions between the gross profitability and value premium, and we show that incorporating a fixed cost of production and its implied operating leverage into our model does not affect the main quantitative results (see Online Appendix 1).

Related Literature. Our paper contributes to the literature on the interaction between the gross profitability and value premium. Novy-Marx (2013) presents the gross profitability premium. Moreover, the gross profitability and value premium become more pronounced after controlling for the book-to-market ratio and gross profitability, respectively, a puzzling pattern that lies at the heart of both anomalies. Feng, Giglio and Xiu (2020) show that the profitability factor, unlike most recently discovered factors, is useful in explaining asset returns, even after accounting for a large set of other factors. Despite mounting empirical evidence, only a few papers offer risk-based theoretical explanations of the profitability premium. Kogan and Papanikolaou (2013) highlight the role of the investment-specific technology (IST) shock as a systematic risk factor priced in the cross section. More recently, Li, Kogan and Zhang (2020) and Li et al. (2020) develop models in which more profitable firms are riskier because they benefit less from the operating hedge offered by intermediate inputs. Our paper differs from these papers in two main aspects. First, their models focus on competitive equilibrium and ignore the heterogeneity of product market competition across industries, which is the focus of our paper. Second, their mechanisms can help explain the within-industry gross profitability premium, but not the cross-industry gross profitability premium. Both the heterogeneity of competition across industries and the cross-industry gross profitability premium are the main focus of our paper.

Our paper also contributes to the growing literature on industry returns. Previous studies have examined the relationship between industry returns and industry information leads and lags (e.g., Moskowitz and Grinblatt, 1999; Croce, Marchuk and Schlag, 2019), demographics (e.g., DellaVigna and Pollet, 2007), industry concentration (e.g., Hou and Robinson, 2006; Ali, Klasa and Yeung, 2009; Giroud and Mueller, 2011; Bustamante and Donangelo, 2017; Corhay, Kung and Schmid, 2020*a*), durability of industry output (e.g., Gomes, Kogan and Yogo, 2009), network concentration and sparsity of an industry (e.g., Herskovic, 2018), expected inflation (e.g., Boudoukh, Richardson and Whitelaw, 1994), consumption risk exposures (e.g., Dittmar and Lundblad, 2017), and political connectedness of an industry (e.g., Belo, Gala and Li, 2013; Addoum and Kumar, 2016). We contribute to the literature by shedding light on the connection between industry returns and two primitive industry characteristics, the market leadership turnover rate and the cash flow loading on expected growth.

Moreover, our paper contributes to the burgeoning literature at the intersection of industrial organization (IO), customer base, and finance. Early contributions to this line of research, including those of Fershtman and Judd (1987), Bolton and Scharfstein (1990), and Aggarwal and Samwick (1999), focus on the interaction between competition and contracting. Recently, many studies have focused on the interaction among competition, customer base, asset pricing, and industry dynamics (e.g., Basak and Pavlova, 2004; Garlappi, 2004; Hou and Robinson, 2006; Novy-Marx, 2007; Aguerrevere, 2009; Carlin, 2009; Gârleanu, Kogan and Panageas, 2012; Carlson et al., 2014; Opp, Parlour and Walden, 2014; Bustamante, 2015; Koijen and Yogo, 2015; Loualiche, 2016; Bustamante and Donangelo, 2017; Corhay, 2017; Garlappi and Song, 2017; Andrei and Carlin, 2018; Chen et al., 2020; Corhay, Kung and Schmid, 2020*a*,*b*; Crouzet and Eberly, 2020; Ao et al., 2021; Babenko, Boguth and Tserlukevich, 2021). Most of these papers focus on one-shot non-collusive Nash equilibria, weheras we consider collusive Nash equilibria. Two exceptions are Chen et al. (2020) and Ao et al. (2021), who study the feedback effect between endogenous competition intensity and distress risk, as well as the distress spillover effects through the endogenous strategic competition mechanism within and across industries. Another exception is Opp, Parlour and Walden (2014) who investigate how competition intensifies endogenously as the discount rate increases. Their model focuses on identical firms producing homogeneous goods within an industry and on industries with different number of firms. By contrast, our model allows firms within an industry to differ and focuses on industries with different turnover rates of market leaders. Moreover, their model is qualitative, whereas ours is quantitative.

More broadly, an increasing number of works study how time-varying discount rates endogenously alter cash flows of firms and thus stock returns by affecting agents' strategic interactions. For example, Garlappi (2004) analyzes the impact of competition on the risk premia of R&D ventures engaged in a multiple-stage patent race. Pástor and Veronesi (2012) develop a novel model with learning to study the asset pricing implications of political uncertainty, in which the firms' investment decisions and the government's policy decision are made simultaneously: the government takes into account the firms' anticipated response, and each firm considers the actions of the other firms as well as the government. The authors investigate price dynamics in the Nash equilibrium. More recently, Pástor and Veronesi (2020) provide an explanation for the "presidential puzzle" by developing a model with endogenous election outcomes driven by voters' time-varying risk aversion. Agents play a simultaneous-move game in deciding which party to elect. Our mechanism relies on the collusive Nash equilibrium and folk theorem (Fudenberg and Maskin, 1986), thus is fundamentally different from those above.

2 Model

In this section, we develop a dynamic asset pricing model to rationalize the gross profitability and value premium across industries, especially to reconcile their important interactions. Our model applies the theoretical machinery on endogenous competition developed by Dou, Ji and Wu (2021) to a quantitative general-equilibrium endowment-based model (i.e., a quantitative Lucas-tree framework). It extends the model of Dou, Ji and Wu (2021) in three important ways. First, in addition to the dispersion of market leadership persistence across industries, we introduce the dispersion of firms' cash flow exposures to expected growth shocks across industries. By calibrating these two primitive industry characteristics and their correlation based on the data, we show that our model connects two industry-level financial metrics, namely, gross profitability and the book-to-market ratio, and industry-level stock return exposures to priced economy-wide shocks through their endogenous interactions with the two primitive industry characteristics. Accordingly, our calibrated model generates the joint pattern of the gross profitability and value premium observed in the data. Second, the two dimensions of industry heterogeneity interact in economically interesting and nontrivial ways in our model as in the data, which cannot be generated by the model of Dou, Ji and Wu (2021). Specifically, the interaction between the gross profitability and value premium and that between competition intensity and the book-to-market ratio in a given industry are missing in the model of Dou, Ji and Wu (2021). Third, rather than exogenously specify a stochastic discount factor (SDF), we build on the habit-formation framework of Campbell and Cochrane (1999), augmented with a predictable component of consumption growth as in Santos and Veronesi (2006, 2010) and Lettau and Wachter (2007). The tight connection between the model-implied SDF and aggregate consumption dynamics further disciplines the model quantitatively, thereby substantially strengthening the quantitative justification of the model's core mechanism, which goes beyond the scope of Dou, Ji and Wu (2021).

2.1 Basic Environment

There is a continuum of atomistic and homogeneous households with access to complete financial markets. The corporate sector comprises a continuum of industries indexed by $i \in \mathcal{I} \equiv [0,1]$ and is owned by these households. Each industry i has n oligopolies (market leaders) and many followers with measure zero.[4] We set $n = 2$ so that each industry is essentially a duopoly. Market leaders are indexed by $j \in \{1,2\}$. We denote a generic firm by ij, referring to firm j in industry i, and its competitor by $i\bar{j}$. Firms produce differentiated

[4]In Online Appendix 3.3, we extend the model to allow a nonzero measure of followers and microfound their competition with leaders. Doing so does not change the main implications of the paper.

perishable goods and set their profit margins strategically to maximize shareholder value.

Although firms optimally choose their own outputs (i.e., firm-level output is highly endogenous), aggregate output is exogenously specified, effectively making our model an endowment economy. A similar top-down modeling approach is adopted by Menzly, Santos and Veronesi (2004) and Santos and Veronesi (2006, 2010). In particular, we denote the aggregate endowment in terms of final goods at time t by E_t and assume that the log aggregate endowment, denoted by $e_t \equiv \ln(E_t)$, evolves as follows:

$$\mathrm{d}e_t = g_t \mathrm{d}t + \sigma_e \mathrm{d}Z_{e,t}, \quad \text{and} \tag{2.1}$$

$$\mathrm{d}g_t = -\kappa(g_t - \overline{g})\mathrm{d}t + \sigma_g \sqrt{g_t - \varsigma}\,\mathrm{d}Z_{g,t}, \tag{2.2}$$

where g_t is the time-varying expected growth rate, $Z_{e,t}$ and $Z_{g,t}$ are two independent standard Brownian motions, and ς is the theoretical lower bound for g_t.

We assume that the growth of the aggregate endowment has a predictable low-frequency component g_t for three reasons. First, this assumption is consistent with the literature on long-run risk, which emphasizes that a small component of consumption growth is persistent and predictable (e.g., Kandel and Stambaugh, 1991; Bansal and Yaron, 2004; Hansen, Heaton and Li, 2008; Bansal, Kiku and Yaron, 2012; Müller and Watson, 2018). Second, fluctuations in the expected growth in demand and output can affect the intensity of competition among industry rivals, as emphasized in the macroeconomics and IO literature (e.g., Haltiwanger and Harrington, 1991; Bagwell and Staiger, 1997; Galeotti and Schiantarelli, 1998; Ivaldi et al., 2007; Nekarda and Ramey, 2013). The primary goal of our model is to investigate the asset pricing implications of endogenous strategic competition; thus, it is necessary for the model to incorporate the low-frequency component g_t in expected growth. Third, the cross-sectional heterogeneity of firms' exposures to fluctuations in expected future cash flow growth is important to generate the value premium (e.g., Cohen, Polk and Vuolteenaho, 2002; Campbell and Vuolteenaho, 2004; Santos and Veronesi, 2010). Incorporating expected growth g_t is a natural way to model long-run cash flow risk in an endowment economy, as in Menzly, Santos and Veronesi (2004), Santos and Veronesi (2006, 2010), and Lettau and

Wachter (2007, 2011), among others.

2.2 Preferences

External Habits, Differentiated Goods, and Customer Bases. There is a representative agent who has a two-level constant elasticity of substitution (CES) preference. In particular, the utility of *final goods consumption*, denoted by C_t, is characterized by

$$U_0 = \mathbb{E}_0 \left[\int_0^\infty u_t(C_t, H_t) \mathrm{d}t \right], \tag{2.3}$$

where the instantaneous utility function is given by

$$u_t(C_t, H_t) = e^{-\rho t} \frac{(C_t - H_t)^{1-\gamma}}{1-\gamma}. \tag{2.4}$$

In Equation (2.4), the variable H_t, γ, and ρ denote an external habit level, the agent's risk aversion, and the subjective discount rate, respectively. The preference falls into the class of external habit formation utilities. Similar to Menzly, Santos and Veronesi (2004) and Santos and Veronesi (2006, 2010), our specification is a continuous-time analog of the preference adopted by Campbell and Cochrane (1999). The external habit level H_t depends on past aggregate consumption. That is, the representative agent derives utility from its consumption relative to the past aggregate consumption path. The external habit level H_t captures a subsistence level of consumption or social externality.

The preference in our model differs from that in the asset pricing literature because the final good C_t is determined by a two-level CES aggregation.[5] First, the demand for C_t is determined through the aggregation of industry composites:

$$C_t = \left[\int_0^1 M_{i,t}^{\frac{1}{\epsilon}} C_{i,t}^{\frac{\epsilon-1}{\epsilon}} \mathrm{d}i \right]^{\frac{\epsilon}{\epsilon-1}}, \tag{2.5}$$

where $C_{i,t}$ is the demand for industry i's composite, and the parameter $\epsilon > 1$ captures

[5]The two-level CES preference is a standard modeling device for agents' preferences and demand system in the international trade literature (e.g., Armington, 1969; Anderson, 1979).

the elasticity of substitution among industry composites. The weight $M_{i,t}$ captures the representative agent's taste for industry i's composite. A higher $M_{i,t}$ reflects a higher utility gain from consuming industry i's composite.

Second, industry i's composite $C_{i,t}$ is further determined by aggregating firm-level differentiated products:

$$C_{i,t} = \left[\sum_{j=1}^{2} \left(\frac{M_{ij,t}}{M_{i,t}} \right)^{\frac{1}{\eta}} C_{ij,t}^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}}, \tag{2.6}$$

where $M_{i,t} = \sum_{j=1}^{2} M_{ij,t}$ also appears in equation (2.5) as the taste for industry i's composite, $C_{ij,t}$ is the demand for firm ij's goods, and the parameter $\eta > 1$ captures the elasticity of substitution among goods in the same industry. The weight $M_{ij,t}/M_{i,t}$ captures the representative agent's taste for firm ij's goods. Consistent with the literature (e.g., Atkeson and Burstein, 2008; Corhay, Kung and Schmid, 2020*a*), we assume that $\eta > \epsilon > 1$, which means that goods within the same industry are more substitutable than those cross different industries.

The taste coefficient $M_{ij,t}$ in the CES aggregator (2.6) can be interpreted as consumers' tendency to continue to buy firm ij's products either because of brand loyalty or consumer inertia (Klemperer, 1995). From a firm's perspective, consumers' taste $M_{ij,t}$ can be seen as firm ij's customer base, which affects the demand for the firm's goods.

External Habit Evolution. The habit level H_t depends on the past consumption process. The effect of habit persistence on risk aversion can be conveniently summarized by the surplus consumption ratio, $S_t \equiv (C_t - H_t)/C_t$, defined as percentage difference between consumption and habit. Following the ideas of Campbell and Cochrane (1999, 2000), Menzly, Santos and Veronesi (2004), Santos and Veronesi (2006, 2010), and Lettau and Wachter (2007), among others, we directly postulate the evolution of $s_t \equiv \ln(S_t)$ as follows:[6]

$$\mathrm{d}s_t = -\phi_s(s_t - \bar{s})\mathrm{d}t + \psi(s_t)\left(\mathrm{d}c_t - \mathbb{E}_t\left[\mathrm{d}c_t\right]\right) + \pi\left(\mathrm{d}g_t - \mathbb{E}_t\left[\mathrm{d}g_t\right]\right), \tag{2.7}$$

[6]Our specification does not lead to the linear habit formation of Constantinides (1990) and Detemple and Zapatero (1991): $H_t = \phi_s \int_{-\infty}^{t} e^{-\phi_s(t-\tau)} C_\tau \mathrm{d}\tau$. However, Li (2015) shows that linear habit persistence has quantitative implications similar to the nonlinear habit persistence of Campbell and Cochrane (1999).

where $c_t \equiv \ln(C_t)$ is the log aggregate consumption of final goods. The sensitivity function $\psi(s_t)$ determines how the habit level is formed from past consumption and is given by

$$\psi(s_t) = \begin{cases} \overline{S}^{-1}\sqrt{1-2(s_t-\overline{s})}-1, & \text{when } s_t \le \hat{s}, \\ 0, & \text{when } s_t > \hat{s}, \end{cases} \tag{2.8}$$

where $\overline{s} \equiv \ln(\overline{S})$, with $\overline{S} = \sigma_e\sqrt{\gamma/\phi_s}$ being the deterministic steady state of S_t as in Campbell and Cochrane (1999), and the threshold is defined as $\hat{s} \equiv \overline{s} + \left(1 - e^{2\overline{s}}\right)/2$. According to equation (2.8), the sensitivity function $\psi(s_t) = 0$ if and only if $s_t \geq \hat{s}$.

In our specification (2.7), the log consumption surplus ratio s_t and the shock to contemporaneous log consumption growth ($\mathrm{d}c_t - \mathbb{E}_t[\mathrm{d}c_t]$) are not perfectly conditionally correlated because of the term $\pi(\mathrm{d}g_t - \mathbb{E}_t[\mathrm{d}g_t])$, which differs from the model of Campbell and Cochrane (1999). This specification is similar in essence to that of Brandt and Wang (2003), who allow s_t to be correlated with other business cycle variables such as inflation, and to that of Lettau and Wachter (2007) and Bekaert, Engstrom and Xing (2009), who introduce shocks to preferences. We set $\pi = \sqrt{2/(\gamma\sigma_g^2)}$ to ensure a constant equilibrium interest rate. A positive π is consistent with the evidence for a negative correlation between expected growth rates and discount rates (e.g., Chen, 2010).

Equilibrium SDF. The marginal utility under preference (2.3) is strictly positive, and thus the aggregate endowment is equal to the aggregate consumption of final goods in equilibrium:

$$E_t = C_t. \tag{2.9}$$

It is straightforward to derive the equilibrium SDF as follows:

$$\Lambda_t = e^{-\rho t}(C_t - H_t)^{-\gamma} = e^{-\rho t}S_t^{-\gamma}C_t^{-\gamma}. \tag{2.10}$$

By plugging the equilibrium condition (2.9) into equation (2.10) and applying Ito's lemma, we derive the dynamics of Λ_t as follows:

$$\frac{\mathrm{d}\Lambda_t}{\Lambda_t} = -r_f \mathrm{d}t - \eta(s_t)\mathrm{d}Z_{e,t} - \varsigma(g_t)\mathrm{d}Z_{g,t}, \tag{2.11}$$

where r_f is the equilibrium risk-free rate and is equal to

$$r_f = \rho + \gamma\varsigma - \frac{\gamma\phi_s}{2}, \tag{2.12}$$

and $\eta(s_t)$ and $\varsigma(g_t)$ are the equilibrium market prices of risk for $Z_{e,t}$ and $Z_{g,t}$, respectively, and are equal to

$$\eta(s_t) \equiv \gamma\sigma_e[1 + \psi(s_t)] \;\text{ and }\; \varsigma(g_t) \equiv \gamma\sigma_g\pi\sqrt{g_t - \varsigma}. \tag{2.13}$$

The market price of risk for $Z_{e,t}$ has the same functional form as in the nonlinear external habit formation model of Campbell and Cochrane (1999). Moreover, the market price of risk for $Z_{g,t}$ is positive and sizeable as in Bansal and Yaron (2004) and Ai and Bansal (2018), where the agent has Epstein-Zin-Weil (EZW) utility with a preference for early resolution of uncertainty.[7] Under relevant calibrations, the market price of risk for $Z_{g,t}$ is much less volatile than that of $Z_{e,t}$. Moreover, consistent with the model solution of Bansal and Yaron (2004), the market price of risk for the long-run growth shock $Z_{g,t}$ is approximately a constant, with $\varsigma(g_t) \approx \gamma\sigma_g\pi\sqrt{\bar{g} - \varsigma}$. Therefore, fluctuations in the discount rate are almost entirely driven by variations in $\eta(s_t)$, which are caused by changes in the log surplus consumption ratio s_t.

Importantly, through the specifications in equations (2.2) and (2.7), our model emphasizes the persistent components of both expected returns and expected future dividend growth, indicating the predictability of returns and dividend growth consistent with empirical findings in the literature (e.g., Binsbergen and Koijen, 2010; Koijen and Nieuwerburgh, 2011).[8]

[7] Dew-Becker (2012) embeds habit formation in EZW utility to generate movements in risk aversion.

[8] Moreover, the model specification in equation (2.7) implies the upward-sloping pattern of the positive

Demand System for Differentiated Products. We derive the representative agent's demand system for differentiated goods from the CES preference in equations (2.5) and (2.6). Let $P_{i,t}$ be the price of industry i's composite. Given $P_{i,t}$ and C_t, we obtain $C_{i,t}$ by solving a standard expenditure minimization problem:

$$C_{i,t} = M_{i,t}\left(\frac{P_{i,t}}{P_t}\right)^{-\epsilon} C_t, \quad \text{with } P_t = \left(\int_0^1 M_{i,t} P_{i,t}^{1-\epsilon} \mathrm{d}i\right)^{\frac{1}{1-\epsilon}}, \tag{2.14}$$

where P_t is the price index of final goods. Without loss of generality, we normalize $P_t \equiv 1$ so that the final goods are the numeraire. Industry-level demand $C_{i,t}$ arises endogenously from the CES aggregation of industry composites in equation (2.5), which provides a micro foundation for the industry-level demand curve exogenously postulated in the model of Dou, Ji and Wu (2021).

Next, given $C_{i,t}$, the demand for firm ij's goods is given by

$$C_{ij,t} = \frac{M_{ij,t}}{M_{i,t}}\left(\frac{P_{ij,t}}{P_{i,t}}\right)^{-\eta} C_{i,t}, \quad \text{with } P_{i,t} = \left(\sum_{j=1}^{2} \frac{M_{ij,t}}{M_{i,t}} P_{ij,t}^{1-\eta}\right)^{\frac{1}{1-\eta}}, \tag{2.15}$$

which arises endogenously from the CES aggregation of differentiated products at the firm level described in equation (2.6). In equation (2.15), the demand for firm ij's goods $C_{ij,t}$ increases with $M_{ij,t}/M_{i,t}$ while the price $P_{ij,t}$ and industry-level demand $C_{i,t}$ are kept constant. Thus, it is natural to consider $M_{ij,t}$ as firm ij's customer base and $M_{i,t}$ as industry i's total customer base. In other words, the customer base determines the demand for firm ij's goods $C_{ij,t}$ and industry i's composite $C_{i,t}$ for given prices (e.g., Gourio and Rudanko, 2014; Dou et al., 2021*b*). Moreover, equation (2.15) implies that firm ij has a greater influence on the price index $P_{i,t}$ when its share $M_{ij,t}/M_{i,t}$ is higher.

By combining equations (2.14) and (2.15), the demand curve faced by firm ij is given by

$$C_{ij,t} \equiv \mathbb{C}_{ij,t}(P_{ij,t}, P_{i\bar{j},t}) = P_{ij,t}^{-\eta} P_{i,t}^{\eta-\epsilon} M_{ij,t} C_t. \tag{2.16}$$

covariance between expected excess returns and subsequent anticipated consumption growth across different horizons, highlighted by Gârleanu, Panageas and Yu (2012).

Equation (2.16) shows that a firm's pricing decision creates externalities to its rival's cash flows through the industry's price index $P_{i,t}$. When $\eta - \epsilon$ is large (i.e., the cross-industry elasticity of substitution is significantly greater than the within-industry elasticity of substitution) and $M_{ij,t}/M_{i\bar{j},t}$ is close to 1 (i.e., market shares are balanced in industry i), the competitor's price $P_{i\bar{j},t}$ has a significant impact on firm ij's demand $C_{ij,t}$ through the effect of $P_{i\bar{j},t}$ on the industry-level price index $P_{i,t}$.

We define effective customer capital by $\widetilde{M}_{ij,t} \equiv M_{ij,t}C_t$. Then, the demand curve faced by firm ij in equation (2.16) can be rewritten as

$$C_{ij,t} \equiv \mathbb{C}_{ij,t}(P_{ij,t}, P_{i\bar{j},t}) = P_{ij,t}^{-\eta} P_{i,t}^{\eta-\epsilon} \widetilde{M}_{ij,t}, \tag{2.17}$$

Clearly, the demand for firm ij's goods $C_{ij,t}$ increases with effective customer capital $\widetilde{M}_{ij,t}$ for given prices.

2.3 Two Dimensions of Heterogeneity Across Industries

In our model, the customer base $M_{ij,t}$ fluctuates over time, driven by Poisson displacement shocks, Brownian shocks, and slow-moving fluctuations in expected growth. We consider two sources of heterogeneity across industries, both of which are reflected in how firms' customer bases evolve in a given industry. One primitive industry characteristic is the displacement rate of industry i's market leaders, denoted by λ_i, and the other is industry i's cash flow exposure to the aggregate expected growth rate, denoted by φ_i. Our analysis focuses on the asset pricing patterns in an economy with heterogeneous λ_i and φ_i. As we show in Section 3.3, the dispersion of these two primitive industry characteristics plays a key role in quantitatively accounting for the joint patterns of the gross profitability and value premium.

We use the Poisson process $N_{i,t}$ with industry-specific intensity λ_i to characterize the occurrence of displacement shocks in industry i. A lower λ_i indicates that market leadership is more resilient in industry i. If displacement occurs over $[t, t + \mathrm{d}t]$ (i.e., if $\mathrm{d}N_{i,t} = 1$), the two current market leaders are displaced by two new market leaders who were previously followers. To generate a non-degenerate distribution of customer bases across firms in the same industry, we assume that the customer bases of the two new market leaders in the industry are "reset" to equal values when displacement occurs over $[t, t + \mathrm{d}t]$. That is, $M_{ij,t} = M_{i\bar{j},t} = M_{i,t}/2$ for the two new market leaders when $\mathrm{d}N_{i,t} = 1$. Because the economy comprises a continuum of industries, industry-specific changes in market leaders are idiosyncratic events for the representative agent. The industry-specific market leadership turnover rate λ_i captures the first source of heterogeneity across industries.

The growth rates of different industries have different loadings on the aggregate expected growth rate g_t. If displacement does not occur over $[t, t + \mathrm{d}t]$ (i.e., if $\mathrm{d}N_{i,t} = 0$), the existing market leaders hold their ground, and firm ij's customer base (i.e., the representative agent's taste for firm ij's products) evolves over $[t, t + \mathrm{d}t]$ according to

$$\frac{\mathrm{d}M_{ij,t}}{M_{ij,t}} = \varphi_i(g_t - \overline{g})\mathrm{d}t + \sigma_M \mathrm{d}W_{ij,t}, \tag{2.18}$$

where the term $\varphi_i(g_t - \overline{g})\mathrm{d}t$ captures the sensitivity of the growth rate of firm ij's cash flow to the aggregate expected growth rate g_t,[9] and the standard Brownian motion $W_{ij,t}$ captures idiosyncratic shocks to firm ij's customer base. The industry-specific loading φ_i captures the second source of heterogeneity across industries.

Integrating both sides of equation (2.18) leads to the evolution equation of the aggregate customer base $M_t \equiv \int_0^1 M_{i,t}\mathrm{d}i$, as follows:

$$\frac{\mathrm{d}M_t}{M_t} = \varphi_M(g_t - \overline{g})\mathrm{d}t, \tag{2.19}$$

where φ_M is the average loading across industries (i.e., φ_M is the mean of the distribution of φ_i across industries). According to equation (2.19), the total customer base M_t follows a stationary process in equilibrium.

[9] For example, Bansal, Dittmar and Lundblad (2005) show that firms' expected future cash flow growth loads differently on aggregate expected growth, and Da and Warachka (2009) provide further empirical support by deriving a measure of expectations regarding firms' future cash flows using analysts' consensus earnings forecasts.

By combining equations (2.1), (2.9), and (2.18), we can derive the evolution equation of firm ij's effective customer capital $\widetilde{M}_{ij,t}$ as follows:

$$\frac{\mathrm{d}\widetilde{M}_{ij,t}}{\widetilde{M}_{ij,t}} = [\overline{g} + (\varphi_i + 1)(g_t - \overline{g})]\,\mathrm{d}t + \sigma_e \mathrm{d}Z_{e,t} + \sigma_M \mathrm{d}W_{ij,t}. \tag{2.20}$$

The heterogeneity of λ_i endogenously generates the gross profitability premium across industries as in Dou, Ji and Wu (2021). The heterogeneity of φ_i generates the value premium as in Da (2009), Santos and Veronesi (2010), Croce, Lettau and Ludvigson (2014), and Li and Zhang (2016), who show that the value premium can be quantitatively explained by the estimated loadings on low-frequency consumption risk — value firms are more exposed to low-frequency consumption risk than growth firms.

2.4 Firms' Optimization Under Strategic Rivalry

Each firm ij faces two choice variables: its product price $P_{ij,t}$ and output $Y_{ij,t}$. Firm ij chooses them simultaneously to maximize its value, considering the externalities of its competitor's price $P_{i\bar{j},t}$ and output $Y_{i\bar{j},t}$.

Our model takes a top-down approach similar to Menzly, Santos and Veronesi (2004) and Santos and Veronesi (2006, 2010). Specifically, in our model, aggregate consumption dynamics are exogenously specified; however, the shares of individual firms' outputs and cash flows as fractions of aggregate consumption are endogenously determined by their customer bases and competition intensity. The firm incurs production costs with intensity $\omega Y_{ij,t}$ to produce a flow of goods with intensity $Y_{ij,t}$ over $[t, t + \mathrm{d}t]$. These production costs are not a deadweight cost incurred by the representative agent. Rather, the production costs incurred by each firm ij can be seen as nonfinancial income received by the representative agent. Linear production technology is commonly adopted in the macroeconomics and IO literature (e.g., Garcia-Macia, Hsieh and Klenow, 2018; Aghion et al., 2019; Bils, Klenow and Ruane, 2020). In Online Appendix 1, we extend our model by incorporating a fixed cost of production and its implied operating leverage as in the model of Carlson, Fisher and

Giammarino (2004), and we show that the main quantitative implications are not changed in the extended model.

Let the firm-level and industry-level profit margins be $\theta_{ij,t} = (P_{ij,t} - \omega)/P_{ij,t}$ and $\theta_{i,t} \equiv (P_{i,t} - \omega)/P_{i,t}$, respectively. We focus on profit margins rather than price levels to increase the transparency of the central economic mechanisms, which does not change the main insights or results of this paper.[10]

The demand function in equation (2.17) can be rewritten in terms of profit margins as follows:

$$\mathbb{C}_{ij,t}(\theta_{ij,t}, \theta_{i\bar{j},t}) = \omega^{-\epsilon}(1 - \theta_{ij,t})^{\eta}(1 - \theta_{i,t})^{\epsilon-\eta}\widetilde{M}_{ij,t}, \tag{2.21}$$

where $\widetilde{M}_{ij,t}$ is firm ij's effective customer capital, defined in equation (2.17).

To maximize its market value, the firm would never produce more than demand $C_{ij,t} = \mathbb{C}_{ij,t}(\theta_{ij,t}, \theta_{i\bar{j},t})$, because goods are immediately perishable and production is costly. Thus, following Gourio and Rudanko (2014), Corhay, Kung and Schmid (2020*a*), and Dou et al. (2021*b*), the demand constraint $Y_{ij,t} \leq \mathbb{C}_{ij,t}(\theta_{ij,t}, \theta_{i\bar{j},t})$ must hold for $Y_{ij,t}$, $Y_{i\bar{j},t}$, $\theta_{ij,t}$, and $\theta_{i\bar{j},t}$ in equilibrium. Moreover, it is optimal for the firm to choose $\theta_{ij,t} > 0$, so the firm will produce up to $\mathbb{C}_{ij,t}(\theta_{ij,t}, \theta_{i\bar{j},t})$ in equilibrium:

$$Y_{ij,t} = \mathbb{C}_{ij,t}(\theta_{ij,t}, \theta_{i\bar{j},t}). \tag{2.22}$$

Because the financial market is frictionless, the firm has no incentive to hoard cash as in Bolton, Chen and Wang (2011), and the operating profit of firm ij is entirely paid out as dividends $D_{ij,t}\mathrm{d}t$ over $[t, t + \mathrm{d}t]$. That is, the firm's dividend flow intensity is given by

$$D_{ij,t} = (P_{ij,t} - \omega)Y_{ij,t} = \theta_{ij,t}(1 - \theta_{ij,t})^{-1}\omega Y_{ij,t}, \quad \text{with} \quad \theta_{ij,t} > 0. \tag{2.23}$$

[10]Detailed discussion of the economic reasons for focusing on profit margins rather than price levels when studying product competition can be found in Dou, Ji and Wu (2021).

Plugging equation (2.22) into equation (2.23) and rearranging terms, we have

$$D_{ij,t} = \Pi_{ij,t}(\theta_{ij,t}, \theta_{i\bar{j},t})\widetilde{M}_{ij,t}, \tag{2.24}$$

where $\Pi_{ij,t}(\theta_{ij,t}, \theta_{i\bar{j},t})$ reflects the profitability per unit of firm ij's effective customer capital $\widetilde{M}_{ij,t}$ and has the following expression:

$$\Pi_{ij,t}(\theta_{ij,t}, \theta_{i\bar{j},t}) = \omega^{1-\epsilon}\theta_{ij,t}(1-\theta_{ij,t})^{\eta-1}(1-\theta_{i,t})^{\epsilon-\eta}. \tag{2.25}$$

Firm ij optimally and strategically chooses $\theta_{ij,t}$ and $Y_{ij,t}$ to maximize its shareholder value as follows:

$$V_{ij,0} = \sup_{\theta_{ij,t}, Y_{ij,t}} \mathbb{E}_0\left[\int_0^{\tau_i} \frac{\Lambda_t}{\Lambda_0} D_{ij,t}\mathrm{d}t\right], \text{ subject to } Y_{ij,t} \leq \mathbb{C}_{ij,t}(\theta_{ij,t}, \theta_{i\bar{j},t}) \tag{2.26}$$

$$= \sup_{\theta_{ij,t}} \mathbb{E}_0\left[\int_0^{\tau_i} \frac{\Lambda_t}{\Lambda_0} \Pi_{ij,t}(\theta_{ij,t}, \theta_{i\bar{j},t})\widetilde{M}_{ij,t}\mathrm{d}t\right], \tag{2.27}$$

where τ_i is the random stopping time at which the market leaders are displaced.

Strategic Competition on Profit Margins. We follow Dou, Ji and Wu (2021) and consider the collusive equilibrium.[11] The two firms in the same industry play a dynamic game, in which the stage games of setting profit margins are played continuously and repeated infinitely with exogenous and endogenous state variables changing over time. There exists a Markov-perfect non-collusive Nash equilibrium, which is the repetition of the one-shot Nash equilibrium. Importantly, there also exist multiple subgame-perfect collusive Nash equilibria in which profit-margin strategies are sustained by conditional punishment strategies. Intuitively, the punishment for deviation is to switch from (tacit) cooperation to a full-blown price war. Theoretically, the punishment for deviation is to switch from the collusive Nash equilibrium to the non-collusive Nash equilibrium, which leads to strictly

[11]Extensive empirical evidence shows that industries are highly concentrated and (tacit) collusion among market leaders is widespread, which has a significant economic impact (e.g., Ivaldi et al., 2007). We discuss empirical evidence of (tacit) collusion in Online Appendix 3.4.

lower profit margins. When deviation occurs at time t, the punishment is implemented with probability $\xi \mathrm{d}t$ over $[t, t+dt]$. The intensity ξ can be viewed as a parameter governing the credibility of the punishment for deviating behavior. A higher ξ reduces the incentive to deviate. Firms' profit-margin strategies depend on both the "payoff-relevant" physical states $x_{i,t} = \{M_{i1,t}, M_{i2,t}, C_t, s_t, g_t\}$ in state space $\mathcal{X}$, as in Maskin and Tirole (1988*a*,*b*), and a set of indicator functions that track whether any firm has previously deviated from a collusive profit-margin agreement, as in Fershtman and Pakes (2000, page 212).[12] The characterization of the non-collusive and collusive equilibria is similar to that of Dou, Ji and Wu (2021) and is discussed in Online Appendix 3.5.

2.5 Discussion of Model Ingredients

Homogeneity. By exploiting the model's homogeneity in terms of industry-level effective customer capital $\widetilde{M}_{i,t} = \widetilde{M}_{i1,t} + \widetilde{M}_{i2,t} = M_{i,t}C_t$ for all firms in each industry $i \in \mathcal{I}$, we can reduce the state space of the model to three state variables to characterize industry i's equilibrium. The three state variables are $M_{i1,t}/M_{i,t}$, s_t, and g_t. In particular, the value function of firm ij in the collusive equilibrium, denoted by $V_{ij}^{C}(M_{i1,t}, M_{i2,t}, C_t, s_t, g_t)$, can be represented by

$$V_{ij}^{C}(M_{i1,t}, M_{i2,t}, C_t, s_t, g_t) \equiv v_{ij}^{C}(M_{i1,t}/M_{i,t}, s_t, g_t)\widetilde{M}_{i,t}. \tag{2.28}$$

We numerically solve the normalized firm values $v_{ij}^{C}(M_{i1,t}/M_{i,t}, s_t, g_t)$ and profit margins $\theta_{ij}^{C}(M_{i1,t}/M_{i,t}, s_t, g_t)$ in the collusive equilibrium.[13] Here, the superscript C denotes the value and policy functions in the collusive equilibrium. Because we focus only on the collusive equilibrium, we omit the superscript C throughout the rest of the paper to ease the notational burden.

Aggregate and Idiosyncratic Shocks. The aggregate state variable g_t determines expected consumption growth. Its evolution is given by equation (2.2), where the aggregate shock $Z_{g,t}$ can be interpreted as the aggregate expected growth shock. The industry characteristic φ_i

[12] For notational simplicity, we omit the indicator states of historical deviations.

[13] See Online Appendix 4 for numerical algorithms.

reflects heterogeneous cash flow loadings on expected growth across industries, as shown in equation (2.18). The expected growth shock $Z_{g,t}$ has a positive market price of risk $\varsigma(g_t)$ in equation (2.11).

The aggregate state variable s_t determines the aggregate discount rate. Its evolution is given by equation (2.7), which incorporates aggregate shocks $Z_{g,t}$ and $Z_{e,t}$, where the latter shock enters equation (2.1) to determine the evolution of the aggregate endowment (consumption). Because $\sigma_e \gg \sigma_g \sqrt{\overline{g} - \varsigma}$, the evolution of s_t is determined mainly by $\mathrm{d}Z_{e,t}$ in equation (2.7). Thus, we can interpret $Z_{e,t}$ as the aggregate discount rate shock; that is, a positive shock, $\mathrm{d}Z_{e,t} > 0$, results in a higher log surplus consumption ratio s_t and a lower discount rate. The shock $Z_{e,t}$ has a positive market price of risk $\eta(s_t)$ in equation (2.11).

The evolution of firm ij's customer base share $M_{ij,t}/M_{i,t}$ is driven by two idiosyncratic shocks, $W_{ij,t}$ and $N_{i,t}$. The idiosyncratic shocks $W_{ij,t}$ ($j = 1, 2$) in equation (2.18) are not crucial for the central mechanisms; however, they are necessary to ensure stationary and non-degenerate industry dynamics in the long run. Intuitively, they can be interpreted as idiosyncratic demand (or "taste") shocks. When the leadership turnover shock occurs ($\mathrm{d}N_{i,t} = 1$), industry i's total customer base $M_{i,t}$ remains unchanged, while the two new market leaders start with the same customer base, i.e., $M_{ij,t} = M_{i\bar{j},t} = M_{i,t}/2$. This model specification ensures that no firm will dominate its industry, even in the long run.

Profitability and Valuation Ratios. Industry-level gross profitability is defined as industry i's gross profits normalized by its assets (i.e., effective customer capital $\widetilde{M}_{i,t} = \widetilde{M}_{i1,t} + \widetilde{M}_{i2,t}$ in the model):

$$GP_{i,t} \equiv \frac{(P_{i,t} - \omega)C_{i,t}}{\widetilde{M}_{i,t}} = \theta_{i,t} \left(\frac{\omega}{1 - \theta_{i,t}} \right)^{1-\epsilon}. \tag{2.29}$$

In equilibrium, the industry-level profit margin $\theta_{i,t}$ is always less than or equal to the monopolistic profit margin $1/\epsilon$, and thus it is straightforward to show that $GP_{i,t}$ is strictly increasing in $\theta_{i,t}$. This implies that the profit margin $\theta_{i,t}$ captures information similar to gross profitability $GP_{i,t}$ in our model. Consistent with empirical patterns, profitability ratios $\Pi_{ij,t}(\theta_{ij,t}, \theta_{i\bar{j},t})$ and $GP_{i,t}$ are stationary in equilibrium.

Similar to endowment-based models for value and growth firms (e.g., Lettau and Wachter, 2007; Tsai and Wachter, 2016), we use the decomposition of assets in place and growth options as a proxy for the book-to-market ratio.[14] One way to approximate the value of assets in place is to focus on the value of existing assets without accounting for any growth effects. Specifically, we consider the value of existing effective customer capital that does not include any growth opportunities, denoted by $\widetilde{M}^{a}_{ij,t}$, and we assume that it decays stochastically over time:

$$\frac{\mathrm{d}\widetilde{M}^{a}_{ij,t}}{\widetilde{M}^{a}_{ij,t}} = \frac{\mathrm{d}\widetilde{M}_{ij,t}}{\widetilde{M}_{ij,t}} - [\overline{g} + (\varphi_i + 1)(g_t - \overline{g})]\mathrm{d}t,$$

where $\overline{g} + (\varphi_i + 1)(g_t - \overline{g})$ is the drift of $\mathrm{d}\widetilde{M}_{ij,t}/\widetilde{M}_{ij,t}$ (see equation (2.20)).

The value of assets in place in the collusive equilibrium, denoted by $V^{a}_{ij,0}$, is the present value of cash flows paid based on the profits generated by effective customer capital $\widetilde{M}^{a}_{ij,t}$ over time. Thus, given the optimal collusive profit margins $\theta_{ij,t}$, it holds that

$$V^{a}_{ij,0} = \mathbb{E}_0\left[\int_0^{\tau_i} \frac{\Lambda_t}{\Lambda_0}\Pi_{ij,t}(\theta_{ij,t}, \theta_{i\bar{j},t})\widetilde{M}^{a}_{ij,t}\mathrm{d}t\right]. \tag{2.30}$$

The value of growth opportunities in the collusive equilibrium is defined by

$$V^{o}_{ij,0} \equiv V_{ij,0} - V^{a}_{ij,0},$$

The difference between the total market value and the value of assets in place gives a generic definition of the value of growth options, which captures the present value of dividends owing to the growth of capital in the future (e.g., Gomes, Kogan and Zhang, 2003, equation (25)).

[14]In production-based models for value and growth firms (e.g., Papanikolaou, 2011; Kogan and Papanikolaou, 2013, 2014; Dou, 2017), the firm's book-to-market ratio is used to measure the relative contribution of assets in place and growth opportunities to firm value, because firms with lower book-to-market ratios are likely to derive most of their value from growth opportunities.

Market Clearing Condition. Last, we discuss the market clearing condition in equation (2.9). From equations (2.5) and (2.6), it follows that $C_t = P_t C_t = \int_0^1 P_{i,t} C_{i,t} \mathrm{d}i$ and $P_{i,t} C_{i,t} = \sum_{j=1}^2 P_{ij,t} C_{ij,t}$. These relationships, together with the product-level market clearing conditions $C_{ij,t} = Y_{ij,t}$ and equation (2.9), lead to

$$E_t = \int_0^1 \left(\sum_{j=1}^{2} P_{ij,t} Y_{ij,t} \right) \mathrm{d}i, \tag{2.31}$$

where the revenue of firm ij is $P_{ij,t} Y_{ij,t} = D_{ij,t} + \omega Y_{ij,t}$; that is, the firm's revenue is the sum of its dividends and production costs (see equation (2.23)). In other words, the aggregate endowment is equal to the total revenue of the corporate sector. Thus, although we exogenously specify the evolution of the aggregate endowment E_t in equation (2.1), the way in which it is split into firm-level revenue $P_{ij,t} Y_{ij,t}$ and further into firm-level dividends $D_{ij,t}$ and production costs $\omega Y_{ij,t}$ is endogenous. This top-down Lucas tree modeling approach is similar in spirit to that adopted by Menzly, Santos and Veronesi (2004) and Santos and Veronesi (2006, 2010); however, firm-level revenue in our model evolves endogenously according an economic mechanism different from that used in these papers.

3 Economic Mechanisms and Quantitative Analyses

In this section, we explain the central economic mechanisms and investigate the quantitative capacity of our model to match the data. Section 3.1 describes the data and empirical measures. Section 3.2 calibrates the model. Section 3.3 discusses the central economic mechanisms. Section 3.4 shows the main asset pricing results. Finally, Section 3.5 quantitatively inspects the model's mechanisms and dissects the effects of the main model ingredients by turning them off one at a time and checking the changes incurred.

3.1 Data and Empirical Measures

We here describe the data and empirical measures used in our quantitative analyses.

Data and Industry Portfolio Returns. We obtain consumption data from the U.S. Bureau of Economic Analysis (BEA) and stock return data from the Center for Research in Security Prices (CRSP). Consumption growth is measured by the log growth rate of per-capita real personal consumption expenditures on non-durable goods and services.

We compute profitability and the book-to-market ratio based on financial data from Compustat. Industry-level gross profitability is constructed as gross profits (revenue minus cost of goods sold) scaled by assets, as defined by Novy-Marx (2013). The industry-level book-to-market ratio is the ratio of book equity to market equity in an industry. Industry-level revenue, cost of goods sold, book assets, book equity, and market equity are the sum of the corresponding firm-level measures for firms in the same industry.

Our model focuses on strategic competition among a few oligopolistic firms whose products are close substitutes. Therefore, we use the SIC4 codes to define industries, following the literature (e.g., Hou and Robinson, 2006; Gomes, Kogan and Yogo, 2009; Frésard, 2010; Giroud and Mueller, 2010, 2011; Bustamante and Donangelo, 2017). We use Compustat segment data to improve the precision of industry classifications (see Online Appendix 2.1).

Our analysis focuses on the gross profitability and value premium across industries. We compute industry-level stock returns as the value-weighted average stock returns of individual firms in a given industry weighted by their market capitalization lagged by 1 month. We exclude financial and utility firms from the analysis and use CRSP delisting returns to adjust for delisted stocks. To ensure that the cross-industry gross profitability and value premium do not simply reflect the firm-level premium, we exclude industries that contain fewer than three firms when computing the cross-industry premium.

Leadership Turnover Rates. We construct an industry-level measure of market leadership turnover rates following Dou, Ji and Wu (2021). In particular, we define market leaders as the top 2 firms ranked by sales in a given industry, which includes both public and private firms. Similar to estimating the probability of corporate events (e.g., Shumway, 2001; Campbell, Hilscher and Szilagyi, 2008), we estimate the leadership turnover rate using a logistic regression model. Specifically, we assume that the marginal probability of market leadership turnover from year t to year $t+1$ follows a logistic regression model given by

$$\mathbb{P}(\mathbb{1}_{turnover,i}^{t\to t+1} = 1) = 1/\left[1 + \exp(-b_0 - b_1 x_{i,t} - \delta_i - \theta_t)\right], \tag{3.1}$$

where $\mathbb{1}_{turnover,i}^{t\to t+1}$ is an indicator that equals 1 if the market leaders of industry i in year $t+1$ are different from those in year t (see Online Appendix 2.1 for more details). The term $x_{i,t}$ is a column vector of explanatory variables whose values are known at the end of year t. Following the IO literature (e.g., Geroski and Toker, 1996; Sutton, 2007; Kato and Honjo, 2009), the vector of explanatory variables $x_{i,t}$ includes industry asset growth rate, industry advertising intensity (i.e., advertising expenses scaled by revenue), industry research and development (R&D) intensity (i.e., R&D expenses scaled by revenue), and the industry-level innovation similarity measure (see Appendix B in Dou, Ji and Wu, 2021). The terms δ_i and θ_t are industry and time fixed effects, respectively.

The leadership turnover measure, denoted by $\hat{\lambda}_{i,t}$, is the predicted probability that one or more existing market leaders are replaced in year $t+1$:

$$\hat{\lambda}_{i,t} = 1/\left[1 + \exp(-\hat{b}_0 - \hat{b}_1 x_{i,t} - \hat{\delta}_i - \hat{\theta}_t)\right], \tag{3.2}$$

where $\hat{b}_0$, $\hat{b}_1$, $\hat{\delta}_i$, and $\hat{\theta}_t$ are estimated using specification (3.1).

We use the average of $\hat{\lambda}_{i,t}$ over time to measure the market leadership turnover rate of industry i, denoted by $\hat{\lambda}_i$. To construct gross-profitability-sorted portfolio-level measures for market leadership turnover rates, we first sort all industries into quintiles (and deciles) according to their gross profitability, and then measure the leadership turnover rate for each portfolio, denoted by $\hat{\lambda}_p$, using the median value of the leadership turnover rates in the industry-year panel within portfolio p. We tabulate $\hat{\lambda}_p$ in panel B of Table 2, which shows that the leadership turnover rate is lower in more profitable industries, consistent with the theoretical and quantitative implications of our model.

Loadings on Expected Growth. We estimate the industry-level cash flow loading on expected growth by running the following time-series regression for each industry i separately using data from 1951 to 2018:

$$\sum_{j=0}^{2} \phi^j ROE_{i,t-j} = \alpha + \varphi_i \sum_{j=0}^{2} \phi^j c_{t-j} + \varepsilon_{i,t}, \tag{3.3}$$

where $\sum_{j=0}^{2} \phi^j ROE_{i,t-j}$ is the accumulated return on equity (ROE) of industry i from year $t-2$ to year t. Following the definition of ROE in Santos and Veronesi (2010), we calculate industry-level ROE in year t as the ratio of industry-level clean-surplus earnings in year t and industry-level book equity in year $t-1$, where clean-surplus earnings in year t are the changes in book equity from year $t-1$ to year t plus dividends in year t. The term $\sum_{j=0}^{2} \phi^j c_{t-j}$ is the accumulated consumption growth from year $t-2$ to year t, a proxy for latent expected future consumption growth used in studies involving low-frequency consumption risk (e.g., Bansal, Dittmar and Lundblad, 2005; Dittmar and Lundblad, 2017). We set ϕ = 0.87 to be consistent with the yearly persistence coefficient of the surplus consumption ratio estimated by Campbell and Cochrane (1999). The coefficient $\hat{\varphi}_i$ estimated using specification (3.3) is the industry-level cash flow loading on expected growth.

To measure the book-to-market-sorted portfolio-level cash flow loading on expected growth, we first sort all industries into quintiles (and deciles) according to their book-to-market ratios, and then follow Santos and Veronesi (2010) to estimate the portfolio-level loading by running the following time series regression for each industry portfolio p separately using data from 1951 to 2018:

$$\sum_{j=0}^{2} \phi^j ROE_{p,t-j} = \alpha + \varphi_p \sum_{j=0}^{2} \phi^j c_{t-j} + \varepsilon_{p,t}, \tag{3.4}$$

where $\sum_{j=0}^{2} \phi^j ROE_{p,t-j}$ is the accumulated ROE of industry portfolio p from year $t-2$ to year t. We measure portfolio-level ROE by computing the value-weighted average industry-level ROE based on industry-level market capitalization lagged by 1 year. For each industry portfolio p, we denote the estimated portfolio-level cash flow loading on expected growth

Table 1: Calibration and parameter choice

Parameter	Symbol	Value	Parameter	Symbol	Value
Panel A: Externally determined parameters					
Risk aversion	γ	2	Volatility of consumption growth	σ_e	0.021
Persistence of surplus ratio	ϕ_s	0.13	Average consumption growth	$\overline{g}$	0.0189
Lower bound of growth	ς	-0.02	Persistence of expected growth	κ	0.3
Volatility of expected growth	σ_g	0.014	Cross-industry elasticity	ϵ	1.1
Volatility of customer base	σ_M	0.01	Within-industry elasticity	η	17
Marginal cost of production	ω	1			
Panel B: Internally calibrated parameters					
Subjective discount factor	ρ	0.185	Punishment rate	ξ	0.037
Copula coefficient of λ_i and φ_i	ϑ	0.159	Range of loadings on g_t	$[\underline{\varphi},\ \overline{\varphi}]$	$[-1,\ 7]$
Range of turnover rates	$[\underline{\lambda},\ \overline{\lambda}]$	$[0.019,\ 0.125]$			

by $\hat{\varphi}_p$. We tabulate $\hat{\varphi}_p$ in panel B of Table 2, which shows that the cash flow loading on expected growth is higher in industries with higher book-to-market ratios, consistent with the theoretical and quantitative implications of our model.

3.2 Calibration

We determine some model parameters based on external information without simulating the model (see panel A of Table 1) while calibrating the remaining model parameters internally by matching important features of the data (see panel B of Table 1).

Externally Determined Parameters. We follow Campbell and Cochrane (1999) and set $\overline{g} = 1.89\%$, $\phi_s = 0.13$, and $\gamma = 2$. We set the volatility of consumption growth at $\sigma_e = 2.1\%$, which is close to the value calibrated by Bansal, Kiku and Yaron (2012). We set the persistence of expected growth at $\kappa = 0.3$, the volatility of expected growth at $\sigma_g = 1.4\%$, and the lower bound of aggregate consumption growth at $\varsigma = -2\%$, ensuring that $2\kappa(\overline{g} - \varsigma) > \sigma_g^2$, i.e., the square-root process (2.2) is well defined. The consumption process implied by these dynamic parameters is consistent with the data in terms of average growth rates, autocorrelations, and variance ratios (see panel A of Table 2).

Importantly, we calibrate our model to capture the characteristics of U.S. industries. We set the within-industry and cross-industry elasticity of substitution at $\eta = 17$ and $\epsilon = 1.1$, respectively. They are broadly consistent with the calibration and estimation in the IO and international trade literature (e.g., Harrigan, 1993; Head and Ries, 2001; Atkeson and Burstein, 2008). We set customer base volatility at a low constant value of $\sigma_M = 0.01$ to capture the idea of a sticky customer base (e.g., Gourio and Rudanko, 2014; Gilchrist et al., 2017). Without loss of generality, the marginal cost of production is normalized to 1, i.e., $\omega = 1$.

Internally Calibrated Parameters. We calibrate the remaining parameters by matching the real risk-free rate, the average gross profit margin, the ranges of λ_i and φ_i across industry portfolios, and the correlation between λ_i and φ_i across industries in the data, which are summarized in Table 2. We run 2,000 independent parallel simulations. For each simulation, we generate a sample of 1,000 industries for 150 years according to the model solution. The first

putting the length of our simulated sample in a range to mimic the data. We then compute the model-implied moments and adjust the parameters until the moments are in line with the data.

We set the subjective discount factor at $\rho = 0.185$ to match the average real risk-free rate between 1948 and 2018. We set the punishment rate at $\xi = 0.037$ to match the average gross profit margin of all industries.

Crucially, we calibrate the bivariate joint distribution of λ_i and φ_i across industries entirely based on the "baseline moments," which only involve the empirical measures of these two primitive industry characteristics without resorting to model-implied information from asset prices. It should be noted that our calibration approach is largely free from the potential overfitting issues caused by inferring the key structural parameters directly from the potentially-misspecified "cross-equation asset pricing restrictions" implied by the model. This is similar in spirit to the recursive estimation procedure advocated by Hansen (2007, 2012) and Chen, Dou and Kogan (2021) for structural asset pricing models. As the main results of this paper, all the asset pricing moments summarized in Table 3, can be interpreted as untargeted moments. In general, the goodness of fit of additional untargeted

Table 2: Moments in the data and model

Panel A: Consumption growth, risk-free rate, and profit margin

Moments	Data	Model	Moments	Data	Model
Average consumption growth (%)	1.89	1.92	VR(2) of consumption growth	1.47	1.39
	[1.51, 2.26]			[1.10, 1.86]	
AC(1) of consumption growth	0.46	0.40	VR(4) of consumption growth	1.89	1.79
	[0.18, 0.70]			[0.85, 3.15]	
AC(4) of consumption growth	0.11	0.09	VR(6) of consumption growth	2.21	2.04
	[−0.20, 0.27]			[0.88, 4.00]	
AC(6) of consumption growth	0.05	0.05	VR(8) of consumption growth	2.43	2.33
	[−0.35, 0.14]			[0.70, 4.72]	
Average real risk-free rate (%)	0.68	0.68	Average gross profit margin (%)	31.39	27.99
	[−0.21, 1.65]			[29.98, 33.00]	

Panel B: Industry characteristics

Portfolios sorted on gross profitability	D1 (low)	Q1	Q2	Q3	Q4	Q5	D10 (high)
$\hat{\lambda}_p$ in the data	0.126	0.121	0.089	0.075	0.070	0.044	0.013
$\hat{\lambda}_p$ in the model	0.125	0.119	0.096	0.072	0.049	0.025	0.019
Portfolios sorted on the book-to-market ratio	D1 (low)	Q1	Q2	Q3	Q4	Q5	D10 (high)
$\hat{\varphi}_p$ in the data	0.62	2.77	3.81	4.75	3.50	5.95	7.98
$\hat{\varphi}_p$ in the model	0.53	0.96	2.79	4.75	6.25	7.55	7.89
Correlation between $\hat{\lambda}_i$ and $\hat{\varphi}_i$ in the data	0.15						
Correlation between λ_i and φ_i in the model	0.15						

Note: This table tabulates the moments in the data and model. The quarterly consumption data are constructed using U.S. BEA data and cover the postwar period from 1948 to 2017. The moments in panel A are computed following Beeler and Campbell (2012), who focus on the sample period from 1948 to 2008 (our moments replicate theirs when we focus on the same sample period). AC(k) of consumption growth refers to the autocorrelation of consumption growth with a k-year lag. VR(k) of consumption growth refers to the variance ratio of consumption growth with a k-year horizon. Real risk-free interest rate is the average difference between the annual returns of 1-month Treasury bills from CRSP and the rate of change in the consumer price index from 1948 to 2018. The estimation of λ, φ, and their correlation in the data is explained in Section 3.1. When constructing the model moments, we simulate a sample of 1,000 industries for 150 years with an 80-year burn-in period. We then compute the model-implied moments as we do for the data. For each moment, the table reports the average value of 2,000 simulations.

moments, particularly the moments that a model attempts to explain, is emphasized as a useful criterion for assessing a model's external validity in the literature on structural estimation.[15]

In particular, we specify the marginal distribution of an industry's leadership turnover rate λ_i and cash flow loading on expected growth φ_i, in stationary equilibrium, as a uniform distribution over the intervals $[\underline{\lambda}, \overline{\lambda}]$ and $[\underline{\varphi}, \overline{\varphi}]$, respectively.[16] We calibrate $\underline{\lambda} = 0.019$ and $\overline{\lambda} = 0.125$ so that the median values of the market leadership turnover rates in portfolios

[15]In the structural estimation literature, there is a long tradition of using one set of moments to estimate a model and another set of untargeted moments to test the model's out-of-sample fit. Recent examples that explicitly stress the importance of untargeted moments include Dou et al. (2021*a*) and Akcigit, Hanley and Serrano-Velarde (2021).

[16]When simulating the model, we discretize the values of λ_i and φ_i in $N = 10$ grids with $\lambda_1 = \underline{\lambda}$, $\lambda_N = \overline{\lambda}$, $\varphi_1 = \underline{\varphi}$, and $\varphi_N = \overline{\varphi}$.

D1 (the 1st decile) and D10 (the 10th decile) of all industries sorted on gross profitability in the model are in line with those of the data (see panel B of Table 2). Similarly, we calibrate $\underline{\varphi} = -1$ and $\overline{\varphi} = 7$ so that the model-implied values of industry portfolios' cash flow loadings on expected growth in portfolios D1 (the 1st decile) and D10 (the 10th decile) of all industries sorted on the book-to-market ratio are in line with those of the data (see panel B of Table 2).[17] Importantly, as an additional "out-of-sample" validation, we report the industry characteristics λ_i and φ_i for different quintile portfolios (i.e., Q1 through Q5) in the data and model for comparison. Overall, the similarity between the data and the model shown in columns Q1 through Q5 of panel B suggests that the distributions and functional forms assumed in the model are reasonable.

We use a flexible parametric method to capture the interdependence between λ_i and φ_i across industries. We denote by $F_1(\lambda_i)$ and $F_2(\varphi_i)$ the marginal cumulative distribution functions of λ_i and φ_i, respectively, and describe the statistical interdependence between λ and φ using the Gaussian copula:

$$C_{\vartheta}^{Gauss}(x_1, x_2) \equiv \Phi_{\vartheta}\left(\Phi^{-1}(x_1), \Phi^{-1}(x_2)\right), \tag{3.5}$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal variable, and $\Phi_{\vartheta}(\cdot,\cdot)$ is the joint cumulative distribution function of a bivariate normal distribution with zero mean, unity variance, and correlation coefficient ϑ.

The parameter ϑ governs the dependence between the two marginal distributions, $F_1(\lambda)$ and $F_2(\varphi)$. A higher ϑ implies that λ_i and φ_i are more positively associated with each other in the cross section of industries. When $\vartheta = 0$, the two variables λ_i and φ_i are independent across industries. We calibrate $\vartheta = 0.159$ to match the correlation between λ_i and φ_i across industries in the data, which is equal to 0.15. Based on the definition of the copula, the joint

[17] The industry's cash flows are proportional to its effective customer capital $\widetilde{M}_{i,t}$, whose cash flow loading on expected growth is $\varphi_i + 1$ (see equation (2.20)).

distribution of (λ_i, φ_i), denoted by $F(\lambda, \varphi)$, can be expressed as follows:

$$F(\lambda, \varphi) = C_{\vartheta}^{Gauss}(F_1(\lambda), F_2(\varphi)). \tag{3.6}$$

The correlation between λ_i and φ_i in the data is statistically significant, with a p-value of 0.024. Intuitively, the positive correlation between λ_i and φ_i at the industry level is consistent with the results of the finance literature. On the one hand, market leaders in industries with higher book-to-market ratios are more subject to displacement threats from followers within the same industries through disruptive innovations (e.g., Gârleanu, Kogan and Panageas, 2012; Kogan and Papanikolaou, 2013, 2014; Kogan et al., 2017; Kogan, Papanikolaou and Stoffman, 2020). That is, industries with higher book-to-market ratios tend to have a higher λ_i. On the other hand, firms in industries with higher book-to-market ratios are more exposed to fluctuations in expected growth in aggregate consumption (e.g., Bansal, Dittmar and Lundblad, 2005; Parker and Julliard, 2005; Hansen, Heaton and Li, 2008; Santos and Veronesi, 2010; Li and Zhang, 2016). That is, industries with higher book-to-market ratios tend to have a higher φ_i. Taken together, these two results suggest that λ_i and φ_i should be positively correlated across industries, which we find based on our estimation in Section 3.1.

3.3 Central Economic Mechanisms

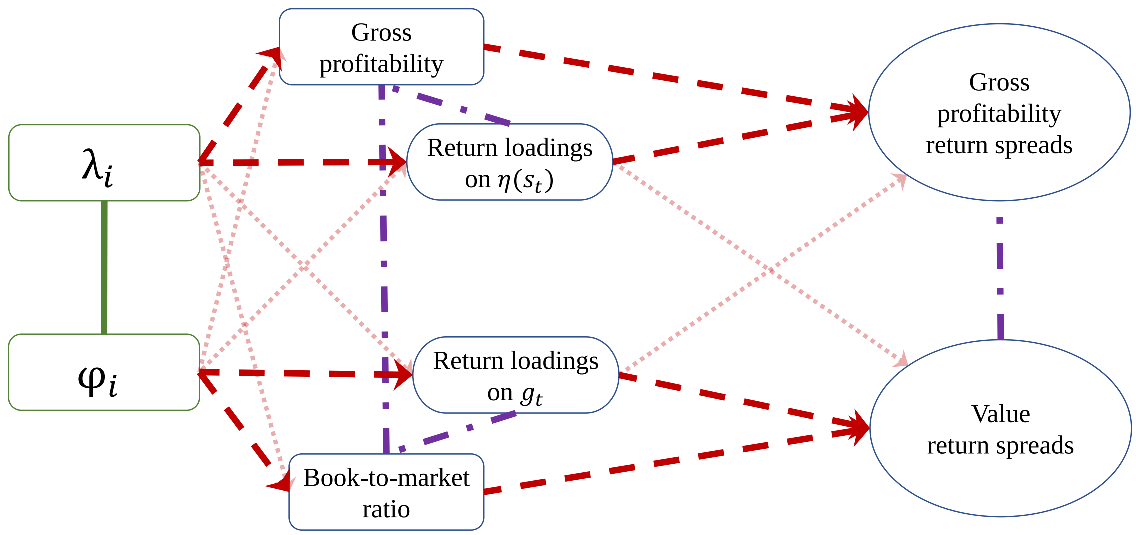
3.3.1 Overview of Challenges and Contributions

Our model combines two economic mechanisms. The first economic mechanism is the one proposed by Dou, Ji and Wu (2021), through which industries with lower leadership turnover rates λ_i have higher profit margins, and both of their profit margins and stock returns are more negatively exposed to fluctuations in the discount rate $\eta(s_t)$. Thus, our model can rationalize the cross-industry gross profitability premium through this mechanism. The second economic mechanism is the one suggested by the empirical findings of Parker and Julliard (2005), Hansen, Heaton and Li (2008), Santos and Veronesi (2010), and Li and Zhang (2016), among others, through which industries with higher cash flow loadings on expected growth φ_i are endogenously associated with higher book-to-market ratios and return exposures to fluctuations in expected growth. Thus, our model also rationalizes the cross-industry value premium.

It may not be surprising that by combining two cross sections (i.e., λ_i and φ_i) and two systematic fluctuations (i.e., changes in $\eta(s_t)$ and g_t), our model is able to quantitatively explain two cross-sectional equity premia (the gross profitability and value premium) simultaneously in a unified framework. It should be noted that our main objective goes substantially beyond separately rationalizing two cross-sectional equity premia within a single model. Motivated by the important insights in the literature,[18] we aim to advance our understanding of the complex and intriguing interactions between the two cross sections, as illustrated in Figure 1. Specifically, by focusing on the interactions between the two cross sections, this paper differs considerably from Dou, Ji and Wu (2021) and contributes to the literature in the following three respects. First, our model shows how gross profitability and book-to-market ratios are endogenously and jointly determined by the two primitive industry characteristics λ_i and φ_i. Second, our model shows how industry-level stock return exposures to $\eta(s_t)$ and g_t are endogenously and jointly determined by the two primitive industry characteristics λ_i and φ_i. Third, based on the two results above, our model demonstrates the interdependence between gross profitability, book-to-market ratios, and industry-level stock return exposures to $\eta(s_t)$ and g_t across industries, thereby rationalizing the interactions between the gross profitability and value premium across industries.

To help visualize the above discussion, Figure 1 illustrates the key variables and their relationships implied by the model and verified by the data. Each of the dashed, dotted, or dash-dotted links represents a relation endogenously implied by the model, and the solid link between the two primitive industry characteristics λ_i and φ_i represents the only exogenously calibrated relation. The arrow links represent the model-implied structural causal relationships between the key variables — the heavy-colored dashed links and the light-colored dotted links with arrows indicate the significant and insignificant endogenous

[18]Influential studies in the literature stress that it is crucial to understand the interactions between the gross profitability and value premium (e.g., Novy-Marx, 2013; Kogan and Papanikolaou, 2013).



Note: This figure illustrates the relationships and interactions among the key variables in the model. The solid link between λ_i and φ_i represents the only exogenously calibrated relationship in our quantitative analysis. The heavy-colored dashed links and the light-colored dotted links with arrows indicate the significant and insignificant structural causal relationships among the key endogenous variables in the model, respectively. The dash-dotted links represent the correlations between the variables that endogenously arise from the model-implied structural causal relationships.

Figure 1: Illustration of the relationships and interactions implied by the model.

structural causal relationships in the model, respectively. The heavy-colored dash-dotted links represent the correlations between the variables that endogenously arise from the model-implied structural causal relationships.

As illustrated in Figure 1, only the correlation and marginal distributional moments of (λ_i, φ_i) across industries — represented by the solid link connecting λ_i and φ_i — are calibrated to match their empirical counterparts, and all other relationships and interactions — represented by the dashed, dotted, and dash-dotted links — are endogenously generated by the economic mechanisms of our model. Put differently, in our quantitative analysis, the correlation and marginal distributional moments of (λ_i, φ_i) across industries are the only targeted moments; all other cross-industry relationships, such as equilibrium asset pricing relationships, are untargeted moments used to test the economic mechanisms.

Overall, the above discussion related to Figure 1 suggests that jointly explaining the gross profitability and value premium across industries, especially their interactions, is a difficult theoretical and quantitative task. We tackle this challenge by showing that a "nearly separating property" holds in the model and verify this property in the data. Specifically, we show that (i) cross-industry dispersions in gross profitability and stock return exposures to the discount rate $\eta(s_t)$, as well as their correlation, are mainly determined by the cross-industry dispersion of λ_i but not of φ_i, and (ii) cross-industry dispersions in book-to-market ratios and stock return exposures to expected growth rate g_t, as well as their correlation, are mainly determined by the cross-industry dispersion of φ_i but not of λ_i. This "nearly separating property" is not obvious ex ante. One of our main contributions is to develop a quantitative model that confirms this property and guides our empirical tests of the key theoretical results in the data. Owing to the "nearly separating property," the interdependence between gross profitability and the book-to-market ratio is endogenously determined by the calibrated correlation between the two primitive industry characteristics (the market leadership turnover rate λ_i and the cash flow loading on expected growth φ_i), and so is the interdependence between the gross profitability and value premium.

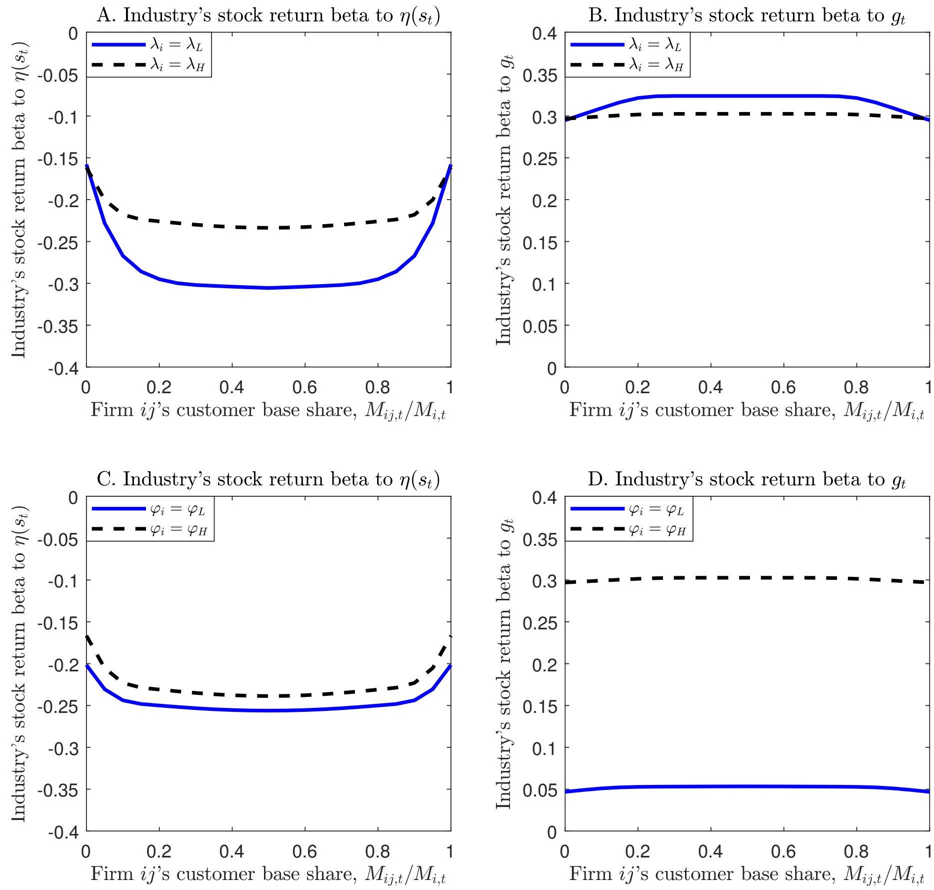
The rest of this subsection is organized as follows. In Section 3.3.2, we first show and discuss the structural causal relationship between the primitive industry characteristic λ_i and an industry's endogenous return exposure to aggregate shocks (i.e., the heavy-colored dashed and the light-colored dotted links with arrows that indicate a strong connection between "λ_i" and "Return loadings on $\eta(s_t)$" and a weak connection between "λ_i" and "Return loadings on g_t," respectively, in Figure 1). In Section 3.3.3, we show and discuss the structural causal relationship between the primitive industry characteristic φ_i and an industry's endogenous return exposure to aggregate shocks (i.e., the heavy-colored dashed and the light-colored dotted links with arrows that indicate a strong connection between "φ_i" and "Return loadings on g_t" and a weak connection between "φ_i" and "Return loadings on $\eta(s_t)$," respectively, in Figure 1). In Section 3.3.4, we show and explain the structural causal relationship between the primitive industry characteristic λ_i and the endogenous profitability/valuation ratios (i.e., the heavy-colored dashed and the light-colored dotted links with arrows that indicate a strong connection between "λ_i" and "Gross profitability" and a weak connection between "λ_i" and "Book-to-market ratio," respectively, in Figure 1). In Section 3.3.5, we show and explain the structural causal relationship between the primitive industry characteristic φ_i and the endogenous profitability/valuation ratios (i.e., the heavy-colored dashed and the light-colored dotted links with arrows that indicate a strong connection between "φ_i" and "Book-to-market ratio" and a weak connection between "φ_i" and "Gross profitability," respectively, in Figure 1). Taken together, these relationships show the "nearly separating property" of the model. Last, in Section 3.3.6, we show that the endogenous correlation between gross profitability and the book-to-market ratio and the endogenous interaction between the gross profitability and value premium (i.e., the heavy-colored dash-dotted links in Figure 1) are consistent with the data when the correlation between λ_i and φ_i is calibrated on the data (i.e., the solid link that connects "λ_i" and "φ_i" in Figure 1). This is because the "nearly separating property" allows us to associate the variation of each sorting variable (gross profitability and the book-to-market ratio) with a distinct cross section of industry primitive characteristics (λ_i and φ_i).

To ensure that the experiments and illustrations are quantitatively meaningful, we adopt the calibrated parameter values from Table 1 in Section 3.2 when presenting the results of the numerical experiments below.

3.3.2 Risk Exposures of Returns in the Cross Section of λ_i

In panels A and B of Figure 2, we illustrate the role of the market leadership turnover rate λ_i in determining an industry's stock return exposures to $\eta(s_t)$ and g_t. In the numerical experiment, we consider two industries with the same φ but different λ_i, with a low value λ_L or a high value λ_H. We set $\varphi = 3$, the median of the calibrated distribution of φ_i in the cross section of industries, and $\lambda_H = 0.07$, the median of the calibrated distribution of λ_i in the cross section of industries, based on the model calibration in Section 3.2. We set $\lambda_L = 0.03$ to capture an industry with a low market leadership turnover rate (i.e., an industry with highly persistent market leadership).

Panels A and B show how industries' stock return betas to fluctuations in $\eta(s_t)$ and g_t, respectively, depend on their market leadership turnover rates λ_i. In particular, panel A shows that the industry with a lower leadership turnover rate λ_L (the blue solid line) has a more negative stock return beta to fluctuations in $\eta(s_t)$ than the industry with a higher



Note: This figure is plotted using the calibrated parameter values in Table 1. Panels A and B plot two industries with different leadership turnover rates (λ_L and λ_H) but the same cash flow loading on expected growth (φ). We set $\lambda_H = 0.07$ and $\varphi = 3$, corresponding to the median values of λ and φ across all industries, respectively. We set $\lambda_L = 0.03$. Panels C and D plot two industries with different cash flow loadings on expected growth (φ_L and φ_H) but the same market leadership turnover rate λ. We set $\varphi_H = 3$ and $\lambda = 0.07$, corresponding to the median values of φ and λ across all industries, respectively. We set $\varphi_L = 0$. Panels A and C plot the industries' stock return betas to fluctuations in $\eta(s_t)$, and we consider a shock that increases $\eta(s_t)$ from its median value ($\eta_L \equiv \eta(s_L)$) to the value of the 95th percentile of the distribution ($\eta_H \equiv \eta(s_H)$). Panels B and D plot the industries' stock return betas to fluctuations in g_t, and we consider a shock that increases g_t from its median value (g_L) to the value of the 95th percentile of the distribution (g_H). Industry-level stock return betas are value-weighted firm-level betas. In panels A and C, industries' stock return betas to fluctuations in $\eta(s_t)$ are given by $\beta_{i,t}^{V,\eta(s)} = \sum_{j=1}^{2} w_{ij,t}\beta_{ij,t}^{V,\eta(s)}$, where $\beta_{ij,t}^{V,\eta(s)} = V_{ij,t}(s_H,\overline{g})/V_{ij,t}(s_L,\overline{g}) - 1$ and $w_{ij,t} = V_{ij,t}(s_L,\overline{g})/\sum_{j'=1}^{2} V_{ij',t}(s_L,\overline{g})$. In panels B and D, industries' stock return betas to fluctuations in g_t are given by $\beta_{i,t}^{V,g} = \sum_{j=1}^{2} w_{ij,t}\beta_{ij,t}^{V,g}$, where $\beta_{ij,t}^{V,g} = V_{ij,t}(\overline{s},g_H)/V_{ij,t}(\overline{s},g_L) - 1$ and $w_{ij,t} = V_{ij,t}(\overline{s},g_L)/\sum_{j'=1}^{2} V_{ij',t}(\overline{s},g_L)$.

Figure 2: Stock return betas in two cross sections.

leadership turnover rate λ_H (the black dashed line). Moreover, panel B shows that the industry with a lower leadership turnover rate λ_L (the blue solid line) has a more positive stock return beta to fluctuations in g_t than the industry with a higher leadership turnover rate λ_H (the black dashed line).

In our model, market leaders' collusion capacity decreases as the discount rate $\eta(s_t)$ increases or expected growth g_t decreases. Intuitively, in the presence of higher discount rates or lower expected growth, market leaders care less about future cooperation and the threat of punishment for deviation; consequently, their collusion capacity is affected by the discount rate $\eta(s_t)$ and expected growth g_t, and thus their profit margins fluctuate endogenously with $\eta(s_t)$ and g_t, which further amplifies the direct effects of $\eta(s_t)$ and g_t on firms' stock returns. In particular, when the discount rate $\eta(s_t)$ increases or expected growth g_t decreases, firm value will decline, not only because of their direct effects, which can be observed from the Gordon valuation formula, but also because of declining profit margins amid intensified industry competition caused by market leaders' lower collusion capacity.

In the cross section, the extent to which firms' collusion capacity is driven by fluctuations in $\eta(s_t)$ and g_t may vary substantially from one industry to another. Specifically, in industries with higher leadership turnover rates λ_i, market leaders are effectively more impatient, and thus the endogenous competition mechanism is less pronounced — market leaders' collusion capacity is less responsive to variations in $\eta(s_t)$ and g_t. Therefore, in industries with higher λ_i, market leaders' profit margins are less exposed to fluctuations in the discount rate $\eta(s_t)$ and expected growth g_t, so their stock return betas to variations in $\eta(s_t)$ and g_t are smaller in magnitude. By showing how competition intensity and profit margins are endogenously driven by fluctuations in expected growth g_t, the above theoretical and quantitative results extend those of Dou, Ji and Wu (2021), who focus on the discount rate shock that drives fluctuations in $\eta(s_t)$.

More importantly, the difference between the solid and dashed lines is very small in panel B but significantly larger in panel A. This stark contrast indicates that as λ_i changes across industries, an industry's the stock return beta to fluctuations in the discount rate $\eta(s_t)$ varies considerably, whereas its stock return beta to fluctuations in expected growth g_t almost stays constant. In fact, the stark contrast between panels A and B highlights an important theoretical and quantitative result of this paper: the endogenous competition channel works for both the discount rate $\eta(s_t)$ and expected growth g_t (i.e., both $\eta(s_t)$ and g_t endogenously drive market leaders' collusion capacity). However, the calibrated model shows that the impact of g_t on stock returns through the endogenous competition channel remains almost unchanged as the market leadership turnover rate λ_i changes across industries, whereas the impact of $\eta(s_t)$ increases significantly from one industry to another as λ_i decreases. These theoretical and quantitative results further strengthen and extend the main result of Dou, Ji and Wu (2021).

3.3.3 Risk Exposures of Returns in the Cross Section of φ_i

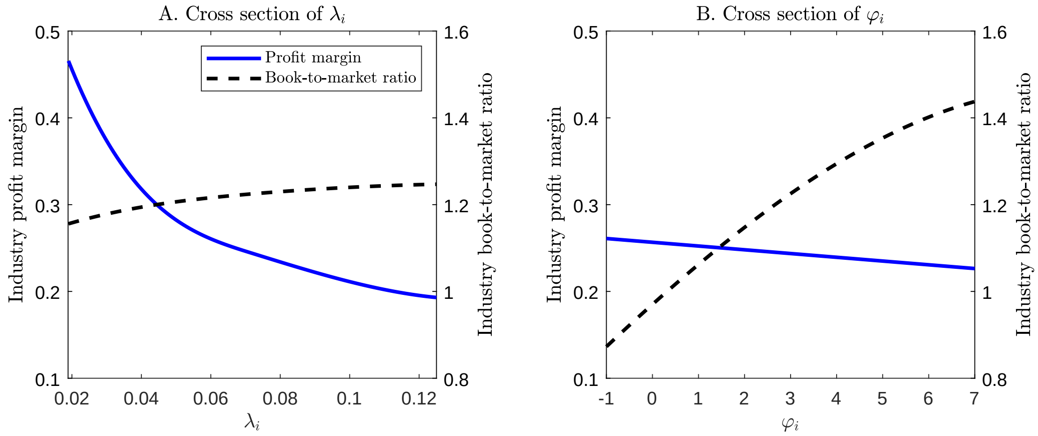
In panels C and D of Figure 2, we illustrate the role of the cash flow loading on expected growth φ_i in determining an industry's stock return exposure to $\eta(s_t)$ and g_t. In the numerical experiment, we consider two industries with the same λ but different φ_i, with a low value φ_L or a high value φ_H. We set $\lambda = 0.07$, the median of the calibrated distribution of λ_i, and $\varphi_H = 3$, the median of the calibrated distribution of φ_i in the cross section of industries, based on the model calibration in Section 3.2. We set $\varphi_L = 0$ to capture an industry with a low cash flow loading on expected growth.

In the cross section, the extent to which firms' stock returns are affected by fluctuations in $\eta(s_t)$ and g_t vary across industries with different cash flow loadings on expected growth φ_i. Panel C shows that the industry with a higher cash flow loading on expected growth φ_H (the black dashed line) has a less negative stock return beta to fluctuations in $\eta(s_t)$ than the industry with a lower cash flow loading on expected growth φ_L (the blue solid line). Indeed, a higher φ_i results in a higher risk premium, which curbs firms' incentive to collude on average. Therefore, the endogenous competition mechanism is indirectly weakened by a higher φ_i because market leaders' collusion capacity and profit margins are less responsive to fluctuations in $\eta(s_t)$. Moreover, panel D shows that the industry with a higher cash flow loading on expected growth φ_H (the black dashed line) has a more positive stock return beta to fluctuations in g_t than the industry with a lower cash flow loading on expected growth φ_L (the blue solid line). The reason is that a firm's cash flow is equal to its profitability per unit of effective customer capital $\Pi_{ij,t}$ multiplied by its effective customer capital $\widetilde{M}_{ij,t}$ (see equation (2.24)). In the industry with a higher cash flow loading on expected growth φ_i, firms' customer base $M_{ij,t}$ responds more strongly to variations in g_t, which in turn results in a more positive loading of $\Pi_{ij,t}$ to fluctuations in g_t through the endogenous competition channel and a more positive loading of $\widetilde{M}_{ij,t}$'s growth to fluctuations in g_t. Thus, a higher φ_i leads to a higher stock return beta to fluctuations in g_t. By showing how industries' stock return betas to fluctuations in $\eta(s_t)$ and g_t are endogenously affected by the primitive industry characteristic φ_i, the above theoretical and quantitative results are beyond the scope of Dou, Ji and Wu (2021), who only focus on one primitive industry characteristic, the market leadership turnover rate λ_i.

More importantly, the difference between the solid and dashed lines is very small in panel C but significantly larger in panel D. This stark contrast indicates that as φ_i changes across industries, an industry's stock return beta to fluctuations in expected growth g_t varies dramatically, whereas its stock return beta to fluctuations in the discount rate $\eta(s_t)$ almost stays constant. Indeed, φ_i directly determines an industry's cash flow loading on expected growth g_t; thus, it strongly affects the dependence of stock returns on expected growth g_t. By contrast, φ_i only indirectly affects how profit margins depend on the discount rate $\eta(s_t)$ through the endogenous competition mechanism; thus, it only weakly affects the dependence of stock returns on the discount rate $\eta(s_t)$. These theoretical and quantitative results are further nontrivial extensions not covered by Dou, Ji and Wu (2021).

3.3.4 Profit Margins and Book-to-Market Ratios in the Cross Section of λ_i

In panel A of Figure 3, the blue solid line shows that industries with a higher λ_i are associated with significantly lower profit margins. As discussed above for panels A and B of Figure 2, the reason is that market leaders in such industries are more impatient because of the higher turnover rate λ_i and thus have less incentive to collude to set high profit margins. By contrast, the black dashed line in panel A of Figure 3 shows that an industry's book-to-market ratio increases with λ_i, but only weakly. This weak relationship across industries is mainly due to the weak association between an industry's exposure to g_t and λ_i under our baseline calibration (see panel B of Figure 2). Intuitively, the difference in industries' book-to-market ratios is mainly caused by industries' heterogeneous exposures to fluctuations in expected growth g_t (see the detailed discussion in Section 3.3.5), which is only weakly associated



Note: This figure is plotted using the calibrated parameter values in Table 1. In panel A, we set the industries' loadings on g_t at $\varphi_i = 3$, the median value of the calibrated distribution of φ in the cross section of industries. In panel B, we set the industries' leadership turnover rates at $\lambda_i = 0.07$, the median value of the calibrated distribution of λ in the cross section of industries. An industry's market and book values correspond to the sum of the market and book values of its two firms, respectively. The market value of firm ij is given by equation (2.27). The book value of firm ij, defined as the value of assets in place, is calculated using equation (2.30).

Figure 3: Profit margins and book-to-market ratios across industries

with λ_i across industries. Although the dispersions of industry-level profit margins and exposures to the discount rate $\eta(s_t)$ are large in the cross section of λ_i (see panel A of Figure 2), they do not lead to large variations in book-to-market ratios. The reason is that profit margins and discount rates affect both the value of assets in place and the value of growth options in roughly the same proportion, having little effect on their ratios.

3.3.5 Profit Margins and Book-to-Market Ratios in the Cross Section of φ_i

In panel B of Figure 3, the blue solid line shows that an industry's profit margin decreases with its cash flow loading on expected growth φ_i, but only slightly. Intuitively, the stock returns of industries with a higher φ_i are more exposed to expected growth g_t, so these industries compensate their shareholders with a higher risk premium. This higher risk premium (i.e., discount rate) effectively makes market leaders more impatient, leading to lower collusive profit margins. However, quantitatively, the variation in profit margins across industries with different φ_i is much smaller than that across industries with different λ_i, mainly because λ_i directly affects the incentive to collude on profit margins, but φ_i only indirectly affects this incentive through the change in the risk premium. The first direct effect of λ_i dominates under the baseline calibration.

By contrast, the black dashed line in panel B of Figure 3 shows that industries with a higher φ_i are associated with significantly higher book-to-market ratios. Intuitively, a higher φ_i directly increases the exposure of an industry's growth to expected growth g_t and thus makes growth options riskier, which reduces the value of these growth options because the future profits created by these options are discounted more aggressively; however, the value of assets in place is not affected by φ_i. Consequently, the book-to-market ratio varies significantly in the cross section of φ_i, which is in sharp contrast to the small change in profit margins in the same cross section of industries. We refer to the effect of φ_i on the book-to-market ratio as the cash flow duration channel similar to that of Lettau and Wachter (2007, 2011), which we provide a discussion in detail in Appendix A.

3.3.6 Interactions Between the Two Cross Sections of λ_i and φ_i

As a key point of our model, we now show that the cross section of industries sorted on gross profitability, as well as the endogenous competition mechanism, interacts with the cross section of industries sorted on book-to-market ratios in economically interesting ways.

Specifically, to rationalize the intriguing correlation between the value return spread and the gross profitability return spread across industries, it is important that the two industry-level sorting variables (i.e., profitability and the book-to-market ratio) have differential sensitivity to changes in the two primitive industry characteristics λ_i and φ_i. Such differential sensitivity arises endogenously in our model, consistent with the discussion of the "nearly separating property". This allows us to associate the variation of each sorting variable with a distinct cross section of industry characteristics. Specifically, sorting industries on gross profitability mainly captures the cross section of industries with different λ_i because profit margins are substantially more sensitive to λ_i than to φ_i across industries (see Figure 3). Industries with lower leadership turnover rates λ_i are associated with higher profit margins (see panel A of Figure 3) and are more negatively exposed to fluctuations in the discount rate $\eta(s_t)$ (see panel A of Figure 2), generating the gross profitability premium. By contrast, sorting industries on the book-to-market ratio mainly captures the cross section of industries with different φ_i because book-to-market ratios are substantially more sensitive to φ_i than to λ_i across industries (see Figure 3). Industries with a higher φ_i are associated with higher book-to-market ratios (see panel B of Figure 3) and are more positively exposed to fluctuations in expected growth g_t (see panel D of Figure 2), generating the value premium.

Furthermore, given the single-sorted portfolios based on the book-to-market ratio, which mainly reflect the cross section of the primitive industry characteristic φ_i, a further sort based on gross profitability within each book-to-market portfolio (i.e., double-sort) almost purely reflects the variation in the primitive characteristic λ_i because the first sort essentially controls for the other industry characteristic φ_i. Given that the two primitive industry characteristics λ_i and φ_i are positively correlated as in the data (see panel B of Table 2), our model implies that the double-sort naturally generates a more pronounced gross profitability premium because λ_i and φ_i have opposite effects on expected equity returns. The same reason explains why the value premium also becomes more pronounced after controlling for gross profitability.

Crucially, the "nearly separating property" ensures that the intriguing correlation between gross profitability and the book-to-market ratio, as well as the interaction between the gross profitability and value premium, ultimately boil down to the correlation between the two cross sections of the primitive industry characteristics λ_i and φ_i, as illustrated in Figure 1. The following hypothetical scenario provides a counterfactual example to illustrate the importance of the correlation between λ_i and φ_i. Suppose that λ_i and φ_i are perfectly negatively correlated across industries. Then, sorting on either gross profitability or book-to-market ratio would capture the same cross section of industries. This would still allow the model to generate both the gross profitability premium and the value premium because industries with a higher λ_i (or equivalently, a lower φ_i) are associated with lower expected returns, lower gross profitability, and lower book-to-market ratios. However, the model-implied interactions between different cross sections are totally off the mark in matching the data. Specifically, the gross profitability premium and the value premium would become less pronounced, rather than more pronounced as in the data, after controlling for the book-to-market ratio and gross profitability, respectively. The main reason is that the cross section of industries captured by the control variable is the same as that captured by the sorting variable. In other words, double-sort analysis essentially sorts twice on the same variable, and the difference among the portfolios created by double sorting is substantially smaller than that among the portfolios created by the first sort.

The above discussion emphasizes the importance of the correlation between λ_i and φ_i to ensure reasonable quantitative performance of the model. As a key result, we show in Section 3.4 that by calibrating the parameter ϑ to match the correlation of 0.15 between the market leadership turnover rate λ_i and the cash flow loading on expected growth φ_i across industries in the data (see panel B of Table 2), our model can quantitatively reproduce the double-sort results observed in the data (see Table 3). To further highlight the key role of the correlation between λ_i and φ_i, we present a systematic discussion of how different correlations between λ_i and φ_i affect the interactions between the industry-level gross profitability and value premium in Table 7 of Section 3.5.

3.4 Asset Pricing Implications

Using the calibrated model, we now ask whether our theory can quantitatively rationalize the industry-level gross profitability and value premium, as well as their interactions. Specifically, in the data, we sort all SIC4 industries into quintiles in June of each year t based on their gross profitability or book-to-market ratios in year $t-1$. Once the portfolios are formed, their monthly industry returns are tracked from July of year t to June of year $t+1$. We compute the portfolio returns by weighting the industry returns in each portfolio with equal weight (EW) and by weighting the industry returns based on the 1-month lagged industry-level market capitalization (VW). In the model, we conduct a similar sorting analysis for the equal- and value-weighted portfolios using simulated industry returns.

Panel A of Table 3 presents the average excess returns of industry portfolios sorted on gross profitability. In both the data and the model, the portfolio consisting of industries with the highest gross profitability (i.e., portfolio Q5) exhibits significantly higher average excess

Table 3: Industry-level gross profitability and value premium in the data and model

	Q1 (low)	Q2	Q3	Q4	Q5 (high)	Q5 − Q1
Panel A: Gross profitability return spreads (sorted on gross profitability)						
Data (EW)	7.23***	7.35***	9.26***	9.16***	10.03***	2.81**
	[2.73]	[3.01]	[3.55]	[3.75]	[4.24]	[2.04]
Data (VW)	6.47***	9.02***	8.77***	9.32***	9.97***	3.50**
	[2.99]	[3.93]	[3.31]	[3.99]	[4.55]	[2.14]
Model (EW)	9.78	9.92	10.25	11.14	12.92	3.14
Model (VW)	9.43	9.56	9.90	10.78	12.60	3.17
Panel B: Gross profitability return spreads after controlling for the book-to-market ratio						
Data (EW)	6.32**	6.28**	9.66***	8.47***	10.11***	3.79***
	[2.43]	[2.21]	[3.91]	[3.14]	[4.20]	[3.22]
Data (VW)	5.50**	6.58**	10.87***	8.35***	9.94***	4.44***
	[2.55]	[2.46]	[4.90]	[3.19]	[4.44]	[2.98]
Model (EW)	9.01	9.56	10.25	11.55	13.64	4.63
Model (VW)	8.99	9.55	10.24	11.55	13.66	4.66
Panel C: Value return spreads (sorted on the book-to-market ratio)						
Data (EW)	5.28**	7.55***	9.04***	9.02***	10.07***	4.79***
	[2.01]	[3.02]	[3.70]	[3.67]	[3.94]	[3.01]
Data (VW)	7.52***	8.13***	8.85***	9.17***	11.07***	3.55*
	[3.13]	[3.77]	[4.25]	[4.28]	[4.97]	[1.87]
Model (EW)	8.08	9.63	11.49	12.24	12.56	4.48
Model (VW)	8.22	9.82	11.70	12.46	12.59	4.37
Panel D: Value return spreads after controlling for gross profitability						
Data (EW)	5.76**	6.93**	7.93***	8.90***	11.43***	5.66***
	[2.19]	[2.53]	[3.23]	[3.28]	[4.53]	[3.93]
Data (VW)	6.65***	7.94***	7.78***	8.11***	11.68***	5.03***
	[2.81]	[3.32]	[3.69]	[3.73]	[5.22]	[2.74]
Model (EW)	7.93	9.21	10.67	12.33	13.87	5.94
Model (VW)	7.97	9.25	10.72	12.37	13.91	5.94

Note: Panel A shows the equal- and value-weighted average excess returns of industry portfolios sorted on gross profitability. In the data, in June of each year t, we sort all industries into quintiles based on their gross profitability in year $t-1$. Once the portfolios are formed, their monthly returns are tracked from July of year t to June of year $t+1$. In the model, we simulate a sample of 1,000 industries for 150 years with an 80-year burn-in period. The 70-year simulated panel places our model sample within the range of the observed data. We then perform portfolio sorting as we do for the data. The excess returns of the portfolios in the model are the average values of 2,000 simulations. In panel B, we perform a double-sort analysis in which we first sort all industries based on their book-to-market ratios, and then we sort the industries in each group into quintiles based on their gross profitability. Panel C shows the equal- and value-weighted average excess returns of the industry portfolios sorted on the book-to-market ratio. In panel D, we perform a double-sort analysis in which we first sort all industries based on their gross profitability and then we sort the industries in each group into quintiles based on their book-to-market ratios. The sample period for the data is from July 1951 to June 2018. We exclude financial firms and utility firms from the analysis. To ensure that the cross-industry gross profitability and value premium do not simply reflect the firm-level premium and avoid potential problematic data points, we exclude all industries with fewer than three firms when computing the cross-industry premium. Newey-West standard errors are estimated with one lag. We annualize the average excess returns by multiplying them by 12. We include t-statistics in brackets. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

returns than that comprising industries with the lowest gross profitability (i.e., portfolio Q1). The equal- and value-weighted gross profitability return spreads (i.e., Q5 − Q1) are 2.81% and 3.50% in the data, respectively, and the model-implied equal- and value-weighted return spreads are 3.14% and 3.17%, respectively. Furthermore, the capital asset pricing model (CAPM) alphas of these quintile portfolios in the data are economically and statistically significant and are reported in Table OA.5 of Online Appendix 2.2.

Panel B of Table 3 presents a double-sort analysis by sorting first on the book-to-market

Table 4: Cross-industry, within-industry, and firm-level gross profitability premium

	Q1 (low)	Q2	Q3	Q4	Q5 (high)	Q5 − Q1
			Panel A: Cross-industry gross profitability return spreads			
EW	7.23***	7.35***	9.26***	9.16***	10.03***	2.81**
	[2.73]	[3.01]	[3.55]	[3.75]	[4.24]	[2.04]
VW	6.47***	9.02***	8.77***	9.32***	9.97***	3.50**
	[2.99]	[3.93]	[3.31]	[3.99]	[4.55]	[2.14]
			Panel B: Within-industry gross profitability return spreads			
VW	5.08**	7.44***	8.23***	8.67***	9.15***	4.08***
	[2.04]	[3.18]	[3.67]	[4.06]	[4.76]	[2.94]
			Panel C: Firm-level gross profitability return spreads			
VW	5.66***	7.21***	7.61***	7.39***	9.89***	4.23***
	[2.59]	[3.56]	[3.74]	[3.59]	[5.14]	[3.11]

Note: Panel A is based on panel A of Table 3. In panel B, we sort all individual firms within each industry (with at least five firms) into quintiles based on their gross profitability lagged by 1 year. In panel C, we sort all firms into quintiles based on their gross profitability lagged by 1 year. The sample period is from July 1951 to June 2018. We exclude financial firms and utility firms from the analysis. Newey-West standard errors are estimated with one lag. We annualize the average excess returns and alphas by multiplying them by 12. We include t-statistics in brackets. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

ratio and then on gross profitability. In both the data and the model, the magnitude of the gross profitability return spread increases after controlling for the book-to-market ratio. The equal- and value-weighted gross profitability return spreads increase to 3.79% and 4.44% in the data, respectively, and to 4.63% and 4.66% in the model, respectively.

Panel C of Table 3 presents the average excess returns of industry portfolios sorted on the book-to-market ratio. In both the data and the model, the industry portfolio with the highest book-to-market ratio (i.e., portfolio Q5) has significantly higher average excess returns than that with the lowest book-to-market ratio (i.e., portfolio Q1). The equal- and value-weighted value return spreads (i.e., Q5 − Q1) are 4.79% and 3.55% in the data, respectively. Correspondingly, the model-implied equal- and value-weighted value return spreads (i.e., Q5 − Q1) are 4.48% and 4.37%, respectively. Moreover, the double-sort analysis in panel D indicates that, after controlling for gross profitability, the equal- and value-weighted value return spreads increase to 5.66% and 5.03% in the data, respectively, and to 5.94% and 5.94% in the model, respectively.

Our paper focuses on the cross-industry gross profitability and value premium, especially their intriguing interactions. Both premia are also prevalent within industries.[19] Tables

[19] In fact, some empirical studies have shown that the gross profitability and value premium cannot be fully explained by industry effects (e.g., Lewellen, 1999; Cohen, Polk and Vuolteenaho, 2003; Novy-Marx, 2013). As a complement, our cross-industry single- and double-sort empirical results show that the relationship between

Table 5: Cross-industry, within-industry and firm-level value premium

	Q1 (low)	Q2	Q3	Q4	Q5 (high)	Q5 − Q1
			Panel A: Cross-industry value return spreads			
EW	5.28**	7.55***	9.04***	9.02***	10.07***	4.79***
	[2.01]	[3.02]	[3.70]	[3.67]	[3.94]	[3.01]
VW	7.52***	8.13***	8.85***	9.17***	11.07***	3.55*
	[3.13]	[3.77]	[4.25]	[4.28]	[4.97]	[1.87]
			Panel B: Within-industry value return spreads			
VW	6.75***	8.07***	8.95***	8.48***	10.63***	3.88***
	[2.97]	[3.80]	[4.32]	[4.11]	[4.71]	[2.69]
			Panel C: Firm-level value return spreads			
VW	6.51***	7.41***	8.46***	9.31***	11.67***	5.16***
	[3.04]	[3.82]	[4.46]	[4.62]	[5.20]	[3.07]

Note: Panel A is based on panel C of Figure 3. In panel B, we sort all individual firms within each industry (with at least five firms) into quintiles based on their book-to-market ratios lagged by 1 year. In panel C, we sort all firms into quintiles based on their book-to-market ratios lagged by 1 year. The sample period is from July 1951 to June 2018. We exclude financial firms and utility firms from the analysis. Newey-West standard errors are estimated with one lag. We annualize the average excess returns and alphas by multiplying them by 12. We include t-statistics in brackets. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

4 and 5 present the cross-industry, within-industry, and firm-level gross profitability and value spreads in average excess returns. The magnitude of the cross-industry premium is comparable to that of the within-industry premium, indicating that cross-industry and within-industry variations are equally important to account for the firm-level premium. The gross profitability spreads and value spreads in CAPM alphas are presented in Tables OA.6 and OA.7 of Online Appendix 2.2, respectively, and the patterns of the CAPM alphas are similar to those of the average excess returns in Tables 4 and 5.

3.5 Quantitative Inspection of the Central Mechanisms

In this section, we conduct counterfactual analyses based on the calibrated model.

3.5.1 Effects of Key Model Ingredients

We first conduct several counterfactual analyses to shed light on the central economic mechanisms and evaluate the quantitative effects of the key ingredients of our model by turning them off one at a time. Column (3) of Table 6 presents the implications of the model for the non-collusive equilibrium. The average profit margin across industries is

the premia at the industry level is similar to that at the firm level documented in the literature.

lower than that of the baseline collusive equilibrium (column (2)) because of the lack of collusion. The growth rate of average net profits is less volatile because profit margins do not vary with $\eta(s_t)$ or g_t in the non-collusive equilibrium, as shown by Dou, Ji and Wu (2021). The equity premium and the volatility of market excess returns are respectively about 1.88% (= 8.79% − 6.91%) and 3.92% (= 18.37% − 14.45%) lower in the non-collusive equilibrium than in the baseline collusive equilibrium, because there is no amplification effect from additional endogenous competition risk. The cross-industry gross profitability premium becomes negligible when the endogenous competition channel is shut down because industry-level profit margins are no longer correlated with λ_i or industry-level risk exposure in the non-collusive equilibrium. The cross-industry value premium is 5.21% in the non-collusive equilibrium, slightly higher than the premium of 4.37% in the collusive equilibrium. The main reason is that the market leadership turnover rate λ_i does not affect industry-level risk exposure in the non-collusive equilibrium, so the positive correlation between φ_i and λ_i does not dampen the value premium. Moreover, the value premium remains roughly unchanged after controlling for gross profitability because gross profitability and the book-to-market ratio are not correlated in the non-collusive equilibrium (see the last row of Table 6).

In columns (4) and (5), we separately quantify the contributions of fluctuations in the discount rate $\eta(s_t)$ and expected growth g_t to generate the gross profitability and value premium across industries. Specifically, to quantify the contribution of fluctuations in the discount rate, in column (4), we set expected growth g_t at its long-run mean $\overline{g}$ while keeping everything else as in the baseline calibration. Comparing columns (2) and (4), we find that fluctuations in the discount rate alone generate a gross profitability premium of 2.99%, whereas the model of baseline calibration with fluctuations in both the discount rate $\eta(s_t)$ and expected growth g_t generates a gross profitability premium of 3.17%. This indicates that the gross profitability premium is mainly due to the time-varying discount rate. Although the value premium is 2.61% in column (4), it is caused by the strong positive correlation between gross profitability and the book-to-market ratio (0.76; see the last row of Table 6), but

Table 6: Inspection of the model mechanisms

	(1) Data	(2) Model	(3) No collusion	(4) $g_t \equiv \overline{g}$	(5) $s_t \equiv \overline{s}$	(6) $\varphi_i \equiv \overline{\varphi}$	(7) $\lambda_i \equiv \overline{\lambda}$
Average gross profit margin (%)	31.39 [29.98, 33.00]	27.71	11.25	28.11	27.00	27.66	23.67
Volatility of the growth rate of net profits (%)	16.22 [11.11, 19.88]	12.01	5.15	8.55	10.53	12.00	9.60
Equity premium ($\mathbb{E}(r - r_f)$, %)	6.68 [2.34, 10.88]	8.79	6.91	6.71	3.69	9.12	8.33
Volatility of market excess returns ($\sigma(r - r_f)$, %)	16.89 [13.21, 19.39]	18.37	14.45	11.65	10.70	19.49	17.17
Sharpe ratio ($\mathbb{E}(r - r_f)/\sigma(r - r_f)$)	0.40 [0.13, 0.77]	0.48	0.48	0.58	0.34	0.47	0.48
Gross profitability premium (sorted on gross profitability $\mathbb{E}(R_{Q5} - R_{Q1})$, %)	3.50 [1.22, 6.06]	3.17	−0.01	2.99	−0.65	4.20	−0.96
Gross profitability premium controlling for the book-to-market ratio (%)	4.44 [1.85, 6.56]	4.66	0.00	1.43	0.67	0.45	0.08
Value premium (sorted on the book-to-market ratio $\mathbb{E}(R_{Q5} - R_{Q1})$, %)	3.55 [1.61, 6.33]	4.37	5.21	2.61	6.77	−4.23	5.93
Value premium controlling for gross profitability (%)	5.03 [3.27, 6.70]	5.94	5.09	0.00	6.77	−0.92	5.78
Correlation between gross profitability and the book-to-market ratio	−0.34 [−0.53, −0.13]	−0.32	0.00	0.76	−0.22	−0.99	−0.03

Note: The sample period is from 1950 to 2018 in the data. R and r are annualized simple returns and log returns, respectively. When constructing the model moments, we simulate a sample of 1,000 industries for 150 years with an 80-year burn-in period. We then compute the model-implied moments as we do for the data. For each moment, the table reports the average value of 2,000 simulations.

not by the dispersion of industry-level cash flow exposures to expected growth g_t. However, the correlation is strongly negative in the data, which explains why, after controlling for gross profitability, the value premium in column (4) decreases to 0, whereas it increases in both the data and the baseline calibration (columns (1) and (2)). Therefore, the results suggest that fluctuations in expected growth g_t play a key role in jointly generating the value premium and the negative correlation between gross profitability and the book-to-market ratio in our model.

In column (5), we quantify the contribution of fluctuations in expected growth by setting the log surplus consumption ratio s_t at its long-run mean $\overline{s}$ while keeping everything else as in the baseline calibration. The cross-industry gross profitability premium is −0.65% in the absence of fluctuations in the discount rate, which clearly contradicts the data. Intuitively, the gross profitability premium should be negative because of the positive correlation of

0.15 between λ_i and φ_i across industries. A lower λ_i leads to a higher profit margin and higher exposure to fluctuations in expected growth g_t through the endogenous competition mechanism. In addition, a lower λ_i is associated with a lower φ_i, leading to a lower risk premium due to less exposure to fluctuations in expected growth g_t. These two channels imply opposite relationships between gross profitability and risk premium across industries, with the latter channel quantitatively dominating under the calibration of column (5). Once we control for the book-to-market ratio (i.e., approximately control for φ_i), the gross profitability premium increases from -0.65% to 0.67%, which mainly reflects the effect of λ_i through the endogenous competition channel described above. The cross-industry value premium is 6.77% in column (5), higher than the premium of 4.37% in column (2), because in the absence of fluctuations in the discount rate, the dispersion of λ_i does not lead to a significant dispersion of risk premia (see panel B of Figure 2) or a significant dispersion of book-to-market ratios (see panel A of Figure 3). Therefore, the positive correlation between λ_i and φ_i, which is 0.15, does not dampen the value premium reflected mainly in the cross section of φ_i alone (see panel D of Figure 2 and panel B of Figure 3).

In columns (6) and (7), we investigate the role played by the two primitive industry characteristics in the cross section. In column (6), we assume that industries have the same loading, $\varphi_i \equiv \overline{\varphi}$, on expected growth g_t, but differ in their leadership turnover rates λ_i. Column (6) shows that the gross profitability premium remains significant at 4.20%, but the value premium is negative, with a value of -4.23%, which strongly contradicts the data. Indeed, when φ_i is the same across industries, sorting on the book-to-market ratio reflects the cross section of λ_i but not that of φ_i. However, panel A of Figure 3 indicates that a higher λ_i is associated with a higher book-to-market ratio but lower gross profitability. Thus, the positive gross profitability premium is naturally coupled with a negative value premium. In column (7), we assume that industries have the same $\lambda_i \equiv \overline{\lambda}$ but different φ_i. Based on the above discussion, unsurprisingly, we find that the value premium remains significant at 5.93%, but the gross profitability premium disappears.

Overall, our counterfactual analyses in columns (4) through (7) clearly suggest that both the time-varying discount rate $\eta(s_t)$ and expected growth g_t, as well as the dispersion of the two primitive industry characteristics λ_i and φ_i across industries, are necessary to simultaneously account for the value premium, the gross profitability premium, the cross-industry correlations between these two premia, and the cross-industry correlation between gross profitability and the book-to-market ratio. The absence of any one of these key model ingredients would prevent the model from matching the data.

3.5.2 Effects of the Correlation Between λ_i and φ_i

We now conduct additional counterfactual analyses to investigate the role of the correlation between the two primitive industry characteristics in explaining the key interaction patterns observed in the data. As discussed above, the theoretical and quantitative model property — "nearly separating property" — arises endogenously. This property ensures that the intriguing correlation between gross profitability and the book-to-market ratio, as well as the interactions between the gross profitability and value premium ultimately, boil down to the correlation between the two cross sections of the primitive industry characteristics λ_i and φ_i. That is, accounting for the complex interactions between the gross profitability and value premium depends on the appropriate calibration of the correlation ϑ between λ_i and φ_i across industries. We now elucidate the role of the key structural parameter ϑ, which governs the correlation between these two primitive industry characteristics.

Column (2) of Table 7 tabulates the baseline calibration where we set $\vartheta = 0.159$ to match the positive correlation of 0.15 between the industry-level leadership turnover rate λ_i and cash flow loading on expected growth φ_i. Using this value of ϑ leads to a negative correlation between gross profitability and the book-to-market ratio (-0.32) close to its empirical counterpart in the data.

In column (3), we set $\vartheta = -0.7$, leading to a significantly negative correlation between λ_i and φ_i of -0.67. As a result, the endogenous correlation between gross profitability and the book-to-market ratio becomes 0.47. The cross-industry gross profitability premium is significantly higher in column (3) than in column (2). This is because a negative correlation

Table 7: Inspection of the correlation between λ_i and φ_i

	(1) Data	(2) Model ($\vartheta = 0.159$)	(3) $\vartheta = -0.7$	(4) $\vartheta = 0$	(5) $\vartheta = 0.7$
Correlation between λ_i and φ_i	0.15 [0.01, 0.28]	0.15	−0.67	0.00	0.67
Correlation between gross profitability and the book-to-market ratio	−0.34 [−0.53, −0.13]	−0.32	0.47	−0.19	−0.75
Gross profitability premium (sorted on gross profitability, $\mathbb{E}(R_{Q5} - R_{Q1})$, %)	3.50 [1.22, 6.06]	3.17	8.49	4.13	0.35
Gross profitability premium controlling for the book-to-market ratio (%)	4.44 [1.85, 6.56]	4.66	4.81	4.92	3.18
Value premium (sorted on the book-to-market ratio, $\mathbb{E}(R_{Q5} - R_{Q1})$, %)	3.55 [1.61, 6.33]	4.37	6.31	4.90	3.37
Value premium controlling for gross profitability (%)	5.03 [3.27, 6.70]	5.94	4.15	6.16	5.39

Note: The sample period is from 1950 to 2018 in the data. All returns are expressed as annualized simple returns. When constructing the model moments, we simulate a sample of 1,000 industries for 150 years with an 80-year burn-in period. We then compute the model-implied moments as we do for the data. For each moment, the table reports the average value of 2,000 simulations.

between λ_i and φ_i indicates that industries with higher gross profitability are associated with lower λ_i and higher φ_i, both of which contribute to a higher risk premium. For a related reason, the value premium is also significantly higher in column (3) than in column (2) because industries with higher book-to-market ratios are associated with a higher φ_i and a lower λ_i. More importantly, the value premium decreases sharply after controlling for gross profitability, and the gross profitability premium decreases significantly after controlling for the book-to-market ratio. Both phenomena clearly contradict our observations in the data and the baseline calibration, as summarized in columns (1) and (2), respectively.

In column (4), we set $\vartheta = 0$, indicating that λ_i and φ_i are independent. In this case, the endogenous correlation of -0.19 between gross profitability and the book-to-market ratio remains negative, but the magnitude is much smaller than the correlation of -0.34 in the data (-0.19 in column (4) but -0.34 in column (1)). It is important to understand why this endogenous correlation remains negative with a value of -0.19, even when $\vartheta = 0$. The main reason is that profit margins decrease monotonically in both λ_i and φ_i, and book-to-market ratios increase monotonically in both λ_i and φ_i (see Figure 3). Thus, within both the cross sections of λ_i and φ_i, profit margins and book-to-market ratios are negatively correlated. This implies that even when λ_i and φ_i are not correlated, the correlation between profitability and the book-to-market ratio remains endogenously negative across industries. With this negative correlation, the model can replicate the pattern whereby the gross profitability and value premium become more pronounced after controlling for the book-to-market ratio and gross profitability, respectively. This result is also reported in column (4) of Table 7. However, the difference in the risk premia with and without controls is clearly smaller than that in our baseline calibration in column (2).

In column (5), we set $\vartheta = 0.7$, which generates a strong negative correlation of -0.75 between gross profitability and the book-to-market ratio; this result is very far from its empirical counterpart and from the implication of the model with the baseline calibration (see columns (1) and (2)). The difference in the risk premia with and without controls is much larger than in our baseline calibration (see column (2)). Specifically, the gross profitability premium increases from 0.35% to 3.18% after controlling for the book-to-market ratio; and the value premium increases from 3.37% to 5.39% after controlling for gross profitability.

4 Conclusion

This paper provides a quantitative explanation of the joint patterns of the gross profitability and value premium across industries. As widely acknowledged in the literature, jointly rationalizing the gross profitability and value premium, especially their interactions, is a difficult task because profitable industries share common characteristics with growth industries, despite their high expected returns. To this end, we develop a novel general equilibrium framework with heterogeneous concentrated industries, consumer inertia, and endogenous strategic competition. Heterogeneity across industries is introduced through cross-sectional differences in two primitive industry characteristics, the market leadership turnover rates and the cash flow loadings on expected growth. Industries with lower market leadership turnover rates are more profitable and responsive to fluctuations in the discount rate through the endogenous competition channel — an increase in the discount rate reduces the present value of future cooperation, causing firms to compete more fiercely for short-run profits by undercutting each other. Meanwhile, industries with higher cash flow loadings on expected growth endogenously have higher book-to-market ratios through the cash flow duration channel and are more responsive to fluctuations in expected growth. The "nearly separating property" of our model ensures that the cross-industry correlation between gross profitability and the book-to-market ratio, as well as the interaction between the gross profitability and value premium ultimately boil down to the correlation parameter of the two primitive industry characteristics. By calibrating the correlation parameter to match its empirical counterpart measured directly in the data, we show that the model has many implications, including the complex interactions between the two cross sections of industries, which is quantitatively consistent with our observations in the data.

References

Addoum, Jawad M, and Alok Kumar. 2016. "Political sentiment and predictable returns." *Review of Financial Studies*, 29(12): 3471–3518.

Aggarwal, Rajesh K., and Andrew A. Samwick. 1999. "Executive compensation, strategic competition, and relative performance evaluation: theory and evidence." *Journal of Finance*, 54(6): 1999–2043.

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J Klenow, and Huiyu Li. 2019. "A theory of falling growth and rising rents." National Bureau of Economic Research Working Paper 26448.

Aguerrevere, Felipe L. 2009. "Real options, product market competition, and asset returns." *Journal of Finance*, 64(2): 957–983.

Ai, Hengjie, and Dana Kiku. 2013. "Growth to value: Option exercise and the cross section of equity returns." *Journal of Financial Economics*, 107(2): 325–349.

Ai, Hengjie, and Ravi Bansal. 2018. "Risk Preferences and the Macroeconomic Announcement Premium." *Econometrica*, 86(4): 1383–1430.

Akcigit, Ufuk, Douglas Hanley, and Nicolas Serrano-Velarde. 2021. "Back to basics: basic research spillovers, innovation policy, and growth." *The Review of Economic Studies*, 88(1): 1–43.

Ali, Ashiq, Sandy Klasa, and Eric Yeung. 2009. "The limitations of industry concentration measures constructed with Compustat data: Implications for finance research." *Review of Financial Studies*, 22(10): 3839–3871.

Anderson, James E. 1979. "A theoretical foundation for the gravity equation." *American Economic Review*, 69(1): 106–116.

Andrei, Daniel, and Bruce I. Carlin. 2018. "The risk of Schumpeterian competition." Working Paper.

Ao, Mingming, Winston Wei Dou, Shane Johnson, and Wei Wu. 2021. "Competition networks, distress propagation, and stock returns." The Wharton School Working Paper.

Armington, Paul S. 1969. "A theory of demand for products distinguished by place of production." *Staff Papers (International Monetary Fund)*, 16(1): 159–178.

Atkeson, Andrew, and Ariel Burstein. 2008. "Pricing-to-market, trade costs, and international relative prices." *American Economic Review*, 98(5): 1998–2031.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2020. "The fall of the labor share and the rise of superstar firms." *Quarterly Journal of Economics*, 135(2): 645âĂŞ709.

Babenko, Ilona, Oliver Boguth, and Yuri Tserlukevich. 2021. "Dynamic competition and expected returns." Arizonal State University Working Paper.

Bagwell, Kyle, and Robert Staiger. 1997. "Collusion over the business cycle." *RAND Journal of Economics*, 28(1): 82–106.

Bansal, Ravi, and Amir Yaron. 2004. "Risks for the long run: A potential resolution of asset pricing puzzles." *Journal of Finance*, 59(4): 1481–1509.

Bansal, Ravi, Dana Kiku, and Amir Yaron. 2012. "An empirical evaluation of the long-run risks model for asset prices." *Critical Finance Review*, 1(1): 183–221.

Bansal, Ravi, Robert F. Dittmar, and Christian T. Lundblad. 2005. "Consumption, dividends, and the cross section of equity returns." *Journal of Finance*, 60(4): 1639–1672.
Basak, Suleyman, and Anna Pavlova. 2004. "Monopoly power and the firm's valuation: a dynamic analysis of short versus long-term policies." *Economic Theory*, 24(3): 503–530.
Beeler, Jason, and John Y. Campbell. 2012. "The long-run risks model and aggregate asset prices: An empirical assessment." *Critical Finance Review*, 1(1): 141–182.
Bekaert, Geert, Eric Engstrom, and Yuhang Xing. 2009. "Risk, uncertainty, and asset prices." *Journal of Financial Economics*, 91(1): 59 – 82.
Belo, Frederico, Vito D Gala, and Jun Li. 2013. "Government spending, political cycles, and the cross section of stock returns." *Journal of Financial Economics*, 107(2): 305–324.
Berk, Jonathan B., Richard C. Green, and Vasant Naik. 1999. "Optimal investment, growth options, and security returns." *Journal of Finance*, 54(5): 1553–1607.
Bils, Mark, Peter J Klenow, and Cian Ruane. 2020. "Misallocation or mismeasurement?" National Bureau of Economic Research Working Paper 26711.
Binsbergen, Jules H. Van, and Ralph S. J. Koijen. 2010. "Predictive regressions: A present-value approach." *Journal of Finance*, 65(4): 1439–1471.
Bolton, Patrick, and David S. Scharfstein. 1990. "A theory of predation based on agency problems in financial contracting." *American Economic Review*, 80(1): 93–106.
Bolton, Patrick, Hui Chen, and Neng Wang. 2011. "A unified theory of Tobin's q, corporate investment, financing, and risk management." *Journal of Finance*, 66(5): 1545–1578.
Boudoukh, Jacob, Matthew Richardson, and Robert F Whitelaw. 1994. "Industry returns and the Fisher effect." *Journal of Finance*, 49(5): 1595–1615.
Brandt, Michael W., and Kevin Q. Wang. 2003. "Time-varying risk aversion and unexpected inflation." *Journal of Monetary Economics*, 50(7): 1457 – 1498.
Bronnenberg, Bart J., Sanjay K. Dhar, and Jean-Pierre H. Dubé. 2009. "Brand history, geography, and the persistence of brand shares." *Journal of Political Economy*, 117(1): 87–115.
Bustamante, M. Cecilia. 2015. "Strategic investment and industry risk dynamics." *Review of Financial Studies*, 28(2): 297–341.
Bustamante, M. Cecilia, and Andres Donangelo. 2017. "Product market competition and industry returns." *Review of Financial Studies*, 30(12): 4216–4266.
Campbell, John, Jens Hilscher, and Jan Szilagyi. 2008. "In search of distress risk." *Journal of Finance*, 63(6): 2899–2939.
Campbell, John Y., and John H. Cochrane. 1999. "By force of habit: A consumption-based explanation of aggregate stock market behavior." *Journal of Political Economy*, 107(2): 205–251.
Campbell, John Y., and John H. Cochrane. 2000. "Explaining the poor performance of consumption-based asset pricing models." *Journal of Finance*, 55(6): 2863–2878.
Campbell, John Y., and Tuomo Vuolteenaho. 2004. "Bad beta, good beta." *American Economic Review*, 94(5): 1249–1275.
Campbell, John Y., Stefano Giglio, Christopher Polk, and Robert Turley. 2018. "An intertemporal CAPM with stochastic volatility." *Journal of Financial Economics*, 128(2): 207–233.
Carlin, Bruce I. 2009. "Strategic price complexity in retail financial markets." *Journal of Financial Economics*, 91(3): 278 – 287.
Carlson, Murray, Adlai Fisher, and Ron Giammarino. 2004. "Corporate investment and asset price dynamics: implications for the cross-section of returns." *Journal of Finance*, 59(6): 2577–2603.
Carlson, Murray, Engelbert J. Dockner, Adlai Fisher, and Ron Giammarino. 2014. "Leaders, followers, and risk dynamics in industry equilibrium." *Journal of Financial and Quantitative Analysis*, 49(2): 321–349.
Chen, Hui. 2010. "Macroeconomic conditions and the puzzles of credit spreads and capital structure." *Journal of Finance*, 65(6): 2171–2212.
Chen, Hui, Winston Wei Dou, and Leonid Kogan. 2021. "Measuring the 'dark matter' in asset pricing models." Working Papers.
Chen, Hui, Winston Wei Dou, Hongye Guo, and Yan Ji. 2020. "Feedback and contagion through distressed competition." Working paper.
Cohen, Randolph B, Christopher Polk, and Tuomo Vuolteenaho. 2003. "The value spread." *Journal of Finance*, 58(2): 609–641.
Cohen, Randolph, Christopher Polk, and Tuomo Vuolteenaho. 2002. "Does risk or mispricing explain the cross-section of stock prices." Working Paper.
Constantinides, George M. 1990. "Habit formation: a resolution of the equity premium puzzle." *Journal of Political Economy*, 98(3): 519–543.
Corhay, Alexandre. 2017. "Industry competition, credit spreads, and levered equity returns." Working Paper.
Corhay, Alexandre, Howard Kung, and Lukas Schmid. 2020*a*. "Competition, markups, and predictable returns." *Review of Financial Studies*, 33(12).

Corhay, Alexandre, Howard Kung, and Lukas Schmid. 2020*b*. "Q: Risk, rents, or growth?" Working Paper.
Croce, Mariano M., Martin Lettau, and Sydney C. Ludvigson. 2014. "Investor Information, Long-Run Risk, and the Term Structure of Equity." *The Review of Financial Studies*, 28(3): 706–742.
Croce, M Max, Tatyana Marchuk, and Christian Schlag. 2019. "The leading premium." National Bureau of Economic Research Working Paper.
Crouzet, Nicolas, and Janice Eberly. 2020. "Rents and intangible capital: A Q+ framework." Northwestern University Working Paper.
Da, Zhi. 2009. "Cash flow, consumption risk, and the cross-section of stock returns." *Journal of Finance*, 64(2): 923–956.
Da, Zhi, and Mitchell Craig Warachka. 2009. "Cashflow risk, systematic earnings revisions, and the cross-section of stock returns." *Journal of Financial Economics*, 94(3): 448–468.
Dechow, Patricia M., Richard Sloan, and Mark T. Soliman. 2004. "Implied Equity Duration: A New Measure of Equity Risk." *Review of Accounting Studies*, 9: 197–228.
DellaVigna, Stefano, and Joshua M Pollet. 2007. "Demographics and industry returns." *American Economic Review*, 97(5): 1667–1702.
Detemple, Jerome B., and Fernando Zapatero. 1991. "Asset prices in an exchange economy with habit formation." *Econometrica*, 59(6): 1633–1657.
Dew-Becker, Ian. 2012. "A model of time-varying risk premia with habits and production."
Dittmar, Robert F., and Christian T. Lundblad. 2017. "Firm characteristics, consumption risk, and firm-level risk exposures." *Journal of Financial Economics*, 125(2): 326–343.
Dou, Winston Wei. 2017. "Embrace or fear uncertainty: Growth options, limited risk sharing, and asset prices." Working Paper.
Dou, Winston Wei, Lucian A. Taylor, Wei Wang, and Wenyu Wang. 2021*a*. "Dissecting bankruptcy frictions." *Journal of Financial Economics*, Forthcoming.
Dou, Winston Wei, Yan Ji, and Wei Wu. 2021. "Competition, profitability, and discount rates." *Journal of Financial Economics*, 140(2): 582–620.
Dou, Winston Wei, Yan Ji, David Reibstein, and Wei Wu. 2021*b*. "Inalienable customer capital, corporate liquidity, and stock returns." *Journal of Finance*, 76(1): 211–265.
Fama, Eugene F., and Kenneth R. French. 1997. "Industry costs of equity." *Journal of Financial Economics*, 43(2): 153 – 193.
Feng, Guanhao, Stefano Giglio, and Dacheng Xiu. 2020. "Taming the Factor Zoo: A Test of New Factors." *The Journal of Finance*, 75(3): 1327–1370.
Fershtman, Chaim, and Ariel Pakes. 2000. "A dynamic oligopoly with collusion and price wars." *RAND Journal of Economics*, 31(2): 207–236.
Fershtman, Chaim, and Kenneth L. Judd. 1987. "Equilibrium incentives in oligopoly." *American Economic Review*, 77(5): 927–940.
Frésard, Laurent. 2010. "Financial strength and product market behavior: The real effects of corporate cash holdings." *Journal of Finance*, 65(3): 1097–1122.
Fudenberg, Drew, and Eric Maskin. 1986. "The Folk theorem in repeated games with discounting or with incomplete information." *Econometrica*, 54(3): 533–54.
Galeotti, Marzio, and Fabio Schiantarelli. 1998. "The cyclicality of markups in a model with adjustment costs: econometric evidence for US industry." *Oxford Bulletin of Economics and Statistics*, 60(2): 121–142.
Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J Klenow. 2018. "How destructive is innovation?" Working Paper.
Garlappi, Lorenzo. 2004. "Risk premia and preemption in R&D ventures." *Journal of Financial and Quantitative Analysis*, 39(4): 843–872.
Garlappi, Lorenzo, and Zhongzhi Song. 2017. "Capital utilization, market power, and the pricing of investment shocks." *Journal of Financial Economics*, 126(3): 447–470.
Gârleanu, Nicolae, Leonid Kogan, and Stavros Panageas. 2012. "Displacement risk and asset returns." *Journal of Financial Economics*, 105(3): 491–510.
Gârleanu, Nicolae, Stavros Panageas, and Jianfeng Yu. 2012. "Technological growth and asset pricing." *Journal of Finance*, 67(4): 1265–1292.
Geroski, P.A., and S. Toker. 1996. "The turnover of market leaders in UK manufacturing industry, 1979-86." *International Journal of Industrial Organization*, 14(2): 141 – 158.
Gilchrist, Simon, Jae Sim, Raphael Schoenle, and Egon Zakrajsek. 2017. "Inflation dynamics during the financial crisis." *American Economic Review*, 107(3): 785–823.
Giroud, Xavier, and Holger M. Mueller. 2010. "Does corporate governance matter in competitive industries?" *Journal of Financial Economics*, 95(3): 312 – 331.
Giroud, Xavier, and Holger M Mueller. 2011. "Corporate governance, product market competition, and equity prices." *Journal of Finance*, 66(2): 563–600.
Gomes, Joao F., Leonid Kogan, and Motohiro Yogo. 2009. "Durability of output and expected stock returns."

Journal of Political Economy, 117(5): 941–986.
Gomes, Joao, Leonid Kogan, and Lu Zhang. 2003. "Equilibrium cross section of returns." *Journal of Political Economy*, 111(4): 693–732.
Gourio, Francois, and Leena Rudanko. 2014. "Customer capital." *Review of Economic Studies*, 81(3): 1102–1136.
Grullon, Gustavo, Yelena Larkin, and Roni Michaely. 2018. "Are U.S. industries becoming more concentrated?" *Review of Finance*, forthcoming.
Gutiérrez, Germán, Callum Jones, and Thomas Philippon. 2019. "Entry Costs and the Macroeconomy." National Bureau of Economic Research, Inc NBER Working Papers.
Haltiwanger, John, and Joseph E. Harrington. 1991. "The impact of cyclical demand movements on collusive behavior." *RAND Journal of Economics*, 22(1): 89–106.
Hansen, Lars Peter. 2007. "Generalized method of moments estimation." In *The New Palgrave Dictionary of Economics. edited by Steven N. Durlauf and Lawrence E. Blume*. Palgrave Macmillan.
Hansen, Lars Peter. 2012. "Proofs for large sample properties of generalized method of moments estimators." *Journal of Econometrics*, 170(2): 325 – 330. Thirtieth Anniversary of Generalized Method of Moments.
Hansen, Lars Peter, John C. Heaton, and Nan Li. 2008. "Consumption strikes back? Measuring long-run risk." *Journal of Political Economy*, 116(2): 260–302.
Harrigan, James. 1993. "OECD imports and trade barriers in 1983." *Journal of International Economics*, 35(1-2): 91–111.
Head, Keith, and John Ries. 2001. "Increasing returns versus national product differentiation as an explanation for the pattern of U.S.-Canada trade." *American Economic Review*, 91(4): 858–876.
Herskovic, Bernard. 2018. "Networks in production: Asset pricing implications." *Journal of Finance*, 73(4): 1785–1818.
Hou, Kewei, and David T. Robinson. 2006. "Industry concentration and average stock returns." *Journal of Finance*, 61(4): 1927–1956.
Ivaldi, Marc, Patrick Rey, Paul Seabright, and Jean Tirole. 2007. "The economics of tacit collusion: Implications for merger control." In *The Political Economy of Antitrust. Contributions to Economic Analysis*, , ed. Vivek Ghosal and Johan Stennek, 217–240. Emerald Group Publishing Limited.
Kandel, Shmuel, and Robert F. Stambaugh. 1991. "Asset returns and intertemporal preferences." *Journal of Monetary Economics*, 27(1): 39–71.
Kato, Masatoshi, and Yuji Honjo. 2009. "The persistence of market leadership: evidence from Japan." *Industrial and Corporate Change*, 18(6): 1107–1133.
Klemperer, Paul. 1995. "Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade." *Review of Economic Studies*, 62(4): 515–539.
Kogan, Leonid, and Dimitris Papanikolaou. 2013. "Firm characteristics and stock returns: The role of investment-specific shocks." *Review of Financial Studies*, 26(11): 2718–2759.
Kogan, Leonid, and Dimitris Papanikolaou. 2014. "Growth opportunities, technology shocks, and asset prices." *Journal of Finance*, 69(2): 675–718.
Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. 2017. "Technological innovation, resource allocation, and growth." *Quarterly Journal of Economics*, 132(2): 665–712.
Kogan, Leonid, Dimitris Papanikolaou, and Noah Stoffman. 2020. "Left behind: Creative destruction, inequality, and the stock market." *Journal of Political Economy*, 128(3): 855–906.
Koijen, Ralph S. J., and Motohiro Yogo. 2015. "The cost of financial frictions for life insurers." *American Economic Review*, 105(1): 445–75.
Koijen, Ralph S.J., and Stijn Van Nieuwerburgh. 2011. "Predictability of returns and cash flows." *Annual Review of Financial Economics*, 3(1): 467–491.
Lettau, Martin, and Jessica A. Wachter. 2007. "Why is long-horizon equity less risky? A duration-based explanation of the value premium." *Journal of Finance*, 62(1): 55–92.
Lettau, Martin, and Jessica Wachter. 2011. "The term structures of equity and interest rates." *Journal of Financial Economics*, 101(1): 90–113.
Lewellen, Jonathan. 1999. "The time-series relations among expected return, risk, and book-to-market." *Journal of Financial Economics*, 54(1): 5–43.
Li, Jun, and Harold H. Zhang. 2016. "Short-run and long-run consumption risks, dividend processes, and asset returns." *Review of Financial Studies*, 30(2): 588–630.
Li, Jun, Leonid Kogan, and Harold H. Zhang. 2020. "Operating hedge and gross profitability premium." Working Paper.
Li, Jun, Leonid Kogan, Harold H. Zhang, and Yifan Zhu. 2020. "Operating leverage and hedging: A tale of two production costs for asset pricing." Working Paper.
Li, Yuming. 2015. "Expected returns and habit persistence." *Review of Financial Studies*, 14(3): 861–899.
Loecker, Jan De, Jan Eeckhout, and Gabriel Unger. 2020. "The rise of market power and the macroeconomic implications." *Quarterly Journal of Economics*, 135(2): 561–644.
Loualiche, Erik. 2016. "Asset pricing with entry and imperfect competition." Working Paper.

Maskin, Eric, and Jean Tirole. 1988*a*. "A theory of dynamic oligopoly, I: Overview and quantity competition with large fixed costs." *Econometrica*, 56(3): 549–69.
Maskin, Eric, and Jean Tirole. 1988*b*. "A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and edgeworth cycles." *Econometrica*, 56(3): 571–599.
Menzly, Lior, Tano Santos, and Pietro Veronesi. 2004. "Understanding predictability." *Journal of Political Economy*, 112(1): 1–47.
Moskowitz, Tobias J, and Mark Grinblatt. 1999. "Do industries explain momentum?" *Journal of Finance*, 54(4): 1249–1290.
Müller, Ulrich K., and Mark W. Watson. 2018. "Long-run covariability." *Econometrica*, 86(3): 775–804.
Nekarda, Christopher J., and Valerie A. Ramey. 2013. "The cyclical behavior of the price-cost markup." Working Paper.
Novy-Marx, Robert. 2007. "An equilibrium model of investment under uncertainty." *Review of Financial Studies*, 20(5): 1461–1502.
Novy-Marx, Robert. 2013. "The other side of value: the gross profitability premium." *Journal of Financial Economics*, 108(1): 1–28.
Opp, Marcus M., Christine A. Parlour, and Johan Walden. 2014. "Markup cycles, dynamic misallocation, and amplification." *Journal of Economic Theory*, 154: 126–161.
Papanikolaou, Dimitris. 2011. "Investment shocks and asset prices." *Journal of Political Economy*, 119(4): 639–685.
Parker, Jonathan A., and Christian Julliard. 2005. "Consumption risk and the cross section of expected returns." *Journal of Political Economy*, 113(1): 185–222.
Pástor, Ľuboš, and Pietro Veronesi. 2006. "Was there a Nasdaq bubble in the late 1990s?" *Journal of Financial Economics*, 81(1): 61–100.
Pástor, Ľuboš, and Pietro Veronesi. 2009. "Technological revolutions and stock prices." *American Economic Review*, 99(4): 1451–1483.
Pástor, Ľuboš, and Pietro Veronesi. 2012. "Uncertainty about Government Policy and Stock Prices." *The Journal of Finance*, 67(4): 1219–1264.
Pástor, Ľuboš, and Pietro Veronesi. 2020. "Political Cycles and Stock Returns." *Journal of Political Economy*, 128(11): 4011–4045.
Santos, Tano, and Pietro Veronesi. 2006. "Labor income and predictable stock returns." *Review of Financial Studies*, 19(1): 1–44.
Santos, Tano, and Pietro Veronesi. 2010. "Habit formation, the cross section of stock returns and the cash-flow risk puzzle." *Journal of Financial Economics*, 98(2): 385–413.
Shumway, Tyler. 2001. "Forecasting bankruptcy more accurately: a simple hazard model." *Journal of Business*, 74(1): 101–124.
Sutton, John. 2007. "Market share dynamics and the "persistence of leadership" debate." *American Economic Review*, 97(1): 222–241.
Tsai, Jerry, and Jessica A. Wachter. 2016. "Rare booms and disasters in a multi-sector endowment economy." *Review of Financial Studies*, 29(5): 1377–1408.
Zhang, Lu. 2005. "The value premium." *Journal of Finance*, 60(1): 67–103.

Appendix

A Discussion of the Cash Flow Duration Channel

Columns (4) and (6) of Table 6 show that dispersion of the cash flow loading on expected growth is crucial for our model to generate the value premium, as well as the interactions between the value and gross profitability premium across industries. An extensive literature has attempted to rationalize the observed value premium. In our model, industries with higher book-to-market ratios tend to have shorter cash flow durations (i.e., their cash flows are weighted more toward the present), and they have higher expected returns as the equity term structure slopes down. This is referred to as the cash flow duration channel for the value premium (e.g., Campbell and Vuolteenaho, 2004; Dechow, Sloan and Soliman, 2004; Lettau and Wachter, 2007, 2011; Santos and Veronesi, 2010). Intuitively, the longer the duration, the longer it takes for shareholders to recover the cash from their investment. Specifically, cash flow duration depends not only on shareholders' expected cash flows over a long time frame, but also on the risk-adjusted rate of return at which these cash flows are discounted. In fact, to stress the importance of the effect of the discount rate in determining cash flow durations, Da (2009), Santos and Veronesi (2010), Croce, Lettau and Ludvigson (2014), and Li and Zhang (2016) point out that not only a firm's temporal cash flow patterns, but also its cash flow's covariance with consumption is important in explaining the observed cross-sectional patterns of stock returns, especially the value premium.[20] Our model also relies on the effect of the discount rate on cash flow durations to generate the value premium. Industries with higher cash flow loadings on expected growth have riskier growth options and thus have higher book-to-market ratios. In addition, they have shorter cash flow duration despite being more sensitive to expected growth shocks, if the market price of risk of expected growth shocks (i.e., $\varsigma(g_t) \approx \gamma\sigma_g\pi\sqrt{\bar{g}-\varsigma}$ in equation (2.11)) is sufficiently large and the cash flow effect sufficiently reinforces the discount effect (i.e., the coefficient π in equation (2.7) is positive and sufficiently large). In fact, the driving force of the value premium in the models of Zhang (2005) and Li and Zhang (2016) is also the reinforcement effect of cash flows on the discount effect. Therefore, as in these papers, the downward-sloping equity term structure implies the value premium in our model.

Moreover, three main approaches are used in the literature to microfound cash flow durations and the equity term structure. The first approach emphasizes the importance of growth options in generating the value premium (e.g., Berk, Green and Naik, 1999; Gomes, Kogan and Zhang, 2003). Similar to these models, growth options are riskier than assets in place in our model. However, unlike these models, the driving force of the value premium in our model is that value firms have shorter cash flow durations than growth firms. The second approach builds on the idea of the creative destruction of innovation and shows that positive IST shocks hurt assets in place while being hedged by growth options (e.g., Papanikolaou, 2011; Gârleanu, Kogan and Panageas, 2012; Kogan and Papanikolaou, 2013, 2014; Kogan et al., 2017; Kogan, Papanikolaou and Stoffman, 2020). Ai and Kiku (2013) also argue that growth options provide a hedge against risks in assets in place because the cost of option exercise is pro-cyclical. A similarity between these models and ours is that value firms have shorter cash flow durations than growth firms. However, unlike these models, growth options are riskier than assets in place in our model. As suggested by Pástor and Veronesi (2006, 2009) and Campbell et al. (2018), uncertainty should play an important role in generating the value premium. In an incomplete-market growth-option framework, Dou (2017) shows that fluctuations in cash flow and investment uncertainty generate a discount effect and an effective IST shock, respectively; thus, the value premium arises in equilibrium. The third approach emphasizes inflexibility caused by the fixed cost of production and the asymmetric adjustment cost of investment (e.g., Carlson, Fisher and Giammarino, 2004; Zhang, 2005), in which the book-to-market ratio is mainly determined by the riskiness of assets in place. Unlike these models, the book-to-market ratio is determined by the riskiness of growth options in our model. In Online Appendix 1, we analyze an extended model and show that incorporating a fixed cost of production and its implied operating leverage into our model does not affect the main quantitative results. Furthermore, in Online Appendix 2.3.2, we provide empirical evidence showing that operating leverage is unlikely to be the channel through which the interaction patterns of the gross profitability and value premium at the industry level arise in the data.

[20] Additional empirical evidence includes Bansal, Dittmar and Lundblad (2005), Parker and Julliard (2005), and Hansen, Heaton and Li (2008), among others.