

Equilibrium Analysis of Asset Prices: Lessons from CIR and APT

Leonid Kogan* Dimitris Papanikolaou†

December 1, 2021

Abstract

Cox, Ingersoll, and Ross (CIR) model proposed a framework for asset pricing in general equilibrium, introducing an explicit description of the macroeconomy into a model of financial markets. The research program started by CIR has been influential and remains highly relevant. In this paper we summarize how the seminal contribution of CIR has seeded our own academic work, with a particular focus on equilibrium analysis of cross-sectional patterns in stock returns.

Keywords: General equilibrium, asset pricing, investment, production, stock returns, yield curve, arbitrage pricing theory

JEL Codes: G10, G12

*NBER and MIT Sloan School of Management, lkogan@mit.edu

†NBER and Kellogg School of Management, d-papanikolaou@kellogg.northwestern.edu

1 Introduction

In 1985, *Econometrica* published two articles by John Cox, Jonathan Ingersoll, and Stephen Ross (CIR). The first was titled “An Intertemporal General Equilibrium Model of Asset Prices” and the second, “A Theory of the Term Structure of Interest Rates.” These papers, which circulated in the working-paper form since the 1970’s, laid out a set of highly influential ideas that have been developed further by multiple strands of academic and applied literature over the subsequent four decades.

The second paper is known to any researcher or practitioner working in the fixed income area: it develops a classical model of risk-free bond prices, the canonical CIR model, which has become an inextricable part of the fixed-income toolkit and led to numerous extensions. The first paper, which is better known in academic circles than among practitioners, is equally foundational. It brings together ideas from the general equilibrium theory and asset pricing theory and develops a combined model of financial markets and real economy. These two papers inspired evolution of thought in the literature over the next four decades.¹

Our own research interests have been heavily influenced by the ideas laid out in the original CIR papers and subsequent developments in the general equilibrium approach to asset pricing. Our goal in this paper is quite modest: we want to share some of our own recent research and to acknowledge the intellectual debt we owe to this pioneering work. We aim to offer a subjective personal perspective rather than a balanced and comprehensive literature review.

¹As of April 2018, Google Scholar reports over twelve thousand of combined citations to these two papers.

1.1 The Cox-Ingersoll-Ross model

At the time when Cox, Ingersoll, and Ross undertook their analysis, the theoretical core of the asset pricing theory included the Capital Asset Pricing Model (CAPM), its dynamic extension (Merton's I-CAPM), and Ross's Arbitrage Pricing Theory (APT). All these models share an important limitation: they take asset returns as their starting point, and derive theoretical predictions on mutual relations between such returns using demand-based argument. For instance, APT starts with a low-dimensional factor structure for asset returns and uses the fact that investors prefer more to less (which in turn implies absence of pure arbitrage opportunities) to show that expected excess returns on all risky assets must be related linearly to their loadings (betas) on the common return factors. CAPM uses a different set of assumptions, including investor preference over the mean and variance of asset returns, to show that expected excess returns are proportional to return covariances with the market portfolio. Both models are silent on the origins of fluctuations in asset prices, that is, the economic mechanism behind the postulated statistical models of returns.

Cox, Ingersoll, and Ross filled the gap in the existing theoretical structure by incorporating the real side of the economy into a dynamic asset-pricing model.² Relative to the existing demand-based pricing models, they explicitly introduced production processes and capital accumulation into an asset-pricing framework. Similarly to the dynamic models like I-CAPM, they also assumed that economic agents are forward-looking and optimizing.

To enhance model tractability, CIR assume that capital accumulation is frictionless

²Recognizing the importance of bridging the asset pricing theories with general equilibrium macroeconomic models, several other researchers pursued this research direction independently around the same time. These include, notably, [Brock \(1982\)](#) and [Prescott and Mehra \(1980\)](#). The latter placed heavier emphasis on methodology and mathematical properties of the framework, while the work of Cox, Ingersoll, and Ross was geared more towards future applications in asset pricing, such as pricing of state-contingent claims and modeling the term structure of interest rates.

and all production processes exhibit constant returns to scale. Specifically, a unit of capital can be used either for investment or for consumption, and can be re-directed among these uses and among all possible production activities without constraints or frictions. If a unit of capital is used for production process j at t , it produces a random amount of output at time $t + 1$, x_{t+1}^j , which is determined by the exogenous production shocks and technological change.³ Capital is subject to depreciation, possibly random, so a single unit at time t at time depreciates to $(1 - \delta_{t+1}^j)$ units at $t + 1$. Because new capital can be re-allocated to consumption at any time at a unit cost, absence of arbitrage implies that, in equilibrium, the price of a unit of capital is always equal to one, and hence *financial* returns on various productive activities are determined solely by productivity shocks. In particular, over the t to $t + 1$ period, financial return on assets of a firm representing ownership of capital invested in production process j equals $x_{t+1}^j - \delta_{t+1}^j$.

The assumption of frictionless investment and constant returns to scale is somewhat limiting (and has been relaxed by the subsequent literature), as it ties firm returns directly to exogenous properties of production processes. In spite of this simplification, the CIR model delivers important endogenous relations: both the risk-free interest rate, as well as the return on the market portfolio, depend on the properties of production processes and investor preferences. To appreciate the importance of this implication, it is useful to relate the CIR model to the Merton's I-CAPM, a seminal dynamic asset pricing model developed in the early 1970s. The I-CAPM relates risk premia on financial assets to their betas on the market portfolio, as well as their betas on the state variables affecting movements in the overall investment opportunity set: the risk-free rate and the Sharpe ratio of the mean-variance efficient portfolio. I-CAPM does not

³The CIR model uses a continuous-time formulation. Nothing is lost qualitatively by describing the logic of the model in discrete time, as we do here.

take a stand on the economic forces driving changes in state variables. By contrast, investment opportunities in the CIR framework are determined by the properties of production processes and investor preferences. Moreover, evolution of investment opportunities over time is closely linked in their model to capital accumulation, technological progress, as well as demand-side factors, such as investors' risk aversion and beliefs.

The ideas proposed by CIR seeded novel theories and guided a voluminous body of empirical work. For instance, subsequent to the original work by CIR, production economy models have been used to model both the aggregate asset markets ([Jermann, 1998](#); [Tallarini, 2000](#); [Boldrin, Christiano, and Fisher, 2001](#); [Kaltenbrunner and Lochstoer, 2010](#)), and, more recently, the cross-sectional patterns in asset returns. In the next section, we describe how these ideas have influenced our own work on the cross-section of stock returns.

2 The cross-section of returns

Theoretical work on the cross-section of returns is inspired to a large degree by several stylized facts. Specifically, prior empirical studies have identified several firm characteristics as robust predictors of future equity returns: examples of such characteristics include firm valuation ratios ([Lakonishok, Shleifer, and Vishny, 1994](#)). These empirical patterns are often inconsistent with the CAPM, in that the characteristics that predict returns are only weakly related to firms' equity market betas ([Fama and French, 1992](#)). Nevertheless, these empirical patterns appear to be largely consistent with versions of Ross' Arbitrage Pricing Theory (APT). That is, empirical multi-factor models that include a small number of long-short portfolios as factors in addition to the market portfolio typically perform better (see, e.g. [Fama and French, 1993](#)). These empirical factor models have proven to be quite successful in pricing a wide variety of trading

strategies that appear to be mispriced by the CAPM, and contributed to increased popularity of “smart beta” or “factor timing” strategies among investors.

To gain a better understanding of the origins and performance of the popular return factors, we would like to know what type of non-diversifiable risk is associated with the firm characteristics that predict returns, or put differently, what type of risk these multi-factor betas capture. A purely statistical model of returns is silent on these questions. Instead, we need economic models with multiple sources of aggregate uncertainty that have heterogeneous impact on the cross-section of asset returns. Such models connect the APT-style return factors with specific sources of economic risk. They provide insights on how structural shocks propagate in the economy, and suggest how to construct trading strategies (factor-mimicking portfolios) that closely track important economic risks.

2.1 Firm characteristics and stock returns

In this paper we focus on a few of the well-known firm characteristics, which have been shown to forecast returns and give rise to systematic return factors. We borrow heavily from the analysis in [Kogan and Papanikolaou \(2013\)](#). Specifically, we consider the relation between average returns and price-earnings ratios (PE, see, e.g., [Rosenberg, Reid, and Lanstein, 1985](#)), market-to-book ratios and Tobin’s q (q , see e.g., [Fama and French, 1992](#); [Lakonishok et al., 1994](#)), investment rates (IK, see e.g., [Titman, Wei, and Xie, 2004](#)), idiosyncratic return volatility (IVOL, see, e.g. [Ang, Hodrick, Xing, and Zhang, 2006](#)), and market betas (MBETA, see e.g., [Black, Jensen, and Scholes, 1972](#); [Frazzini and Pedersen, 2014](#)).

Table 1 summarizes return spreads between the top and the bottom decile portfolios of firms sorted on each of the characteristics above. There is a declining pattern of average returns across the characteristics-sorted portfolios; the difference in average

returns ranges from -2% for the MBETA-decile portfolios, to -8.9% P/E-decile portfolios. Further, for each one of the five characteristics covered here, the top decile portfolios have higher market betas than the bottom decile portfolios. As a result, the CAPM severely misprices these portfolios: CAPM alphas of the long-short decile portfolios (top minus bottom decile) range from -5.7% for the portfolios sorted on market beta to -10.8% for the portfolios sorted on idiosyncratic volatility.

Table 1: Characteristics, comovement, and risk premia

	10 minus 1 decile portfolios				
	Tobin's q	I/K	P/E	MBETA	IVOL
$E(R_p)$ (%)	-8.79 (-3.26)	-4.94 (-1.42)	-8.91 (-3.00)	-1.97 (-0.52)	-5.86 (-0.95)
σ (%)	20.75	24.86	19.72	26.47	37.05
α (%)	-10.27 (-3.64)	-8.04 (-2.41)	-10.15 (-3.51)	-5.71 (-1.87)	-10.78 (-1.83)
β^{mkt}	0.29 (1.66)	0.62 (2.91)	0.25 (1.62)	0.75 (4.26)	0.98 (3.97)
R^2 (%)	6.55	20.13	5.11	25.93	22.85

This table contains moments of the difference in returns between the top and the bottom decile portfolios of firms sorted on Tobin's q, past investment (I/K), price-earnings ratio (P/E), market beta (MBETA), and idiosyncratic volatility (IVOL). We compute the mean and volatility of each time series, and the results of CAPM regressions. See main text for variable definitions. All decile portfolios are value-weighted and rebalanced annually; portfolios based on accounting variables (q , I/K, and P/E) are rebalanced every June; portfolios based on firm moments (MBETA and IVOL) are rebalanced at the beginning of every calendar year. Estimation is performed at annual frequency. We compute t -statistics (in parentheses) using the Newey-West procedure with three lags. The sample period is 1964–2008 and excludes firms producing investment goods, financial firms (SIC 6000–6799) and utilities (SIC 4900–4949). See [Kogan and Papanikolaou \(2013\)](#) for more details.

Importantly, a key piece of the puzzle is that firms with similar characteristics comove with each other. As we see in Table 1, long-short decile portfolios are very volatile: their annual standard deviations range from 19.7% to 37.1%. These portfolios

are highly diversified, hence their high volatility indicates that they are exposed to systematic risk factors. Yet, their market exposure does not fully account for their volatility: the R^2 of the market model regression ranges from 6.6% for the q -sort to 25.9% for the MBETA-sort. Thus, characteristic-sorted portfolios are exposed to systematic risk factors uncorrelated with the market portfolio. This pattern is particularly striking for the MBETA-sort. Grouping firms based on their market exposures results in portfolios that have systematic risk that is *not* spanned by the market portfolio. Our results imply that a firm’s market beta is cross-sectionally correlated with the firm’s exposure to other systematic risk factors, uncorrelated with the market.

Empirically, the above patterns in returns are well described by two-factor APT model, in which the first factor is the market portfolio and the second factor accounts for a substantial fraction of the residual comovement among all characteristic-sorted portfolios. To show this, we first regress each of the decile portfolios within each of the five cross-sections on the market return and construct the residual, $\tilde{R}_p^e = R_p^e - \beta_{pm}R_m^e$. We then normalize all of the residual return series to unit standard deviation, and extract the first principal component from the residual return series within each cross-section individually and in a pooled cross-section of the extreme decile portfolios (1, 2, 9, and 10) sorted on q , I/K, E/P, MBETA, and IVOL.

As Table 2 shows, there is substantial comovement *among* the I/K-, P/E-, q -, MBETA-, and IVOL-factors. The correlation between each individual factor and the first principal component of the pooled cross-section ranges from 46.8% for the idiosyncratic volatility factor to 92% for the investment factor. Hence, not only do high-I/K firms comove more with other high-I/K firms, but these firms also comove with high P/E, q , MBETA, and to some extent high-IVOL firms.⁴ The magnitude

⁴This result is not driven by the same stocks being ranked similarly using each of the above

of this common source of return variation is substantial: the normalized eigenvalue associated with the first principal component from the pooled cross-section of twenty portfolios is 33.2%. This pattern suggests that the missing risk factor is largely common across the five characteristic-based sorts. The first principal component $PC1$ extracted from the pooled cross-section of extreme decile portfolios is essentially the average of long-short portfolios across the I/K, P/E, q , MBETA, and IVOL sorts.

Table 2: Return comovement across characteristics

	Cross-sections					Eigenvalues
	I/K	P/E	IVOL	MBETA	q	$\lambda_1 / \sum \lambda_i$
I/K						35.5
P/E	72.2					35.6
IVOL	40.7	22.4				51.5
MBETA	81.0	63.9	53.7			41.0
q	80.6	62.1	11.4	60.6		35.9
ALL (PC1)	92.0	77.9	46.8	89.7	74.2	33.2

This table shows return comovement across the five decile portfolio sorts, on I/K, P/E, IVOL, MBETA, and q . We extract the first principal component in each of the five cross-sections from the return residuals from a market model regression, $\tilde{R}_p^e = R_p^e - \beta_{pm} R_m^e$. We fix the sign of the principal components so that they load positively on the top decile portfolios. We repeat the same steps in the pooled cross-section of the extreme decile portfolios (1, 2, 9, and 10) sorted on q , I/K, P/E, MBETA, and IVOL. We extract principal components using the correlation matrix. See main text and notes to Table 1 for more details.

We find that a two-factor APT model—which includes the market return ($R_{mkt,t}$) and the return on the $PC1$ portfolio ($R_{pc1,t}$)—prices returns on each set of the characteristic-sorted portfolios. In particular, we estimate

$$R_{p,t} - r_{f,t-1} = \alpha_p + \beta_{mkt,p} (R_{mkt,t} - r_{f,t}) + \beta_{z,p} R_{pc1,t} + \varepsilon_{p,t}, \quad (1)$$

characteristics — correlations among portfolio assignments using various characteristics are low; pairwise correlations range from 11.3% to 38.1%, with the exception of the Tobin’s q and price-earnings ratio pair, which is 63%.

where $R_{p,t}$ is the portfolio return and $r_{f,t-1}$ is the risk-free rate. As we show in Table 3, the two-factor model (1) captures the spreads in average returns in the cross-sections sorted by q , P/E, I/K, IVOL, and MBETA. The estimates of α are small across the decile portfolios and the Gibbons, Ross, and Shanken (1989) (GRS) test p -values are greater than 10% in each cross-section.

Table 3: The empirical factor model

	10 minus 1 decile portfolios				
	Tobin's q	I/K	P/E	MBETA	IVOL
α (%)	-2.18	1.32	-3.17	3.21	-4.44
	(-1.00)	(0.62)	(-1.41)	(1.27)	(-0.94)
β^{mkt}	0.29	0.62	0.25	0.75	0.98
	(3.83)	(5.16)	(2.66)	(6.94)	(4.24)
β^{PC1}	0.63	0.73	0.54	0.69	0.49
	(7.32)	(10.01)	(5.83)	(8.16)	(3.75)
R^2 (%)	66.13	75.75	54.24	70.47	34.32
GRS	0.647	0.668	1.303	0.947	0.688
	(0.91)	(0.89)	(0.21)	(0.57)	(0.88)

This table summarizes the performance of the two-factor model that includes the market portfolio and the first principal component of the pooled cross-section ($PC1$, normalized to standard deviation of 10%) of q , I/K, P/E, MBETA, and IVOL portfolios. We report the alpha, market beta, and $PC1$ -beta of the top minus bottom decile portfolio of firms sorted on each characteristic, as well as the test statistic from the GRS test that the proposed factor model prices the decile portfolios in each cross-section. We compute the asymptotic covariance matrix of the estimates of regression coefficients using the Newey-West procedure with three lags. We report the t -statistics on the regression coefficients and the p -values of the GRS test in parentheses. See the main text and notes to Tables 1 for more details.

Examining the data more closely, we observe that the systematic return factor $PC1$ earns a negative risk premium. The annualized information ratio of the $PC1$ portfolio is equal to -0.51. APT logic implies that portfolios loading positively on this factor must earn relatively low average returns.

The above facts are reminiscent of the findings of Fama and French (1996), who

document common sources of variation among firms with similar book-to-market and price-earnings ratios. Our analysis thus far is silent on the nature of economic risks giving rise to the PC1 factor, and does not reveal why firms with different investment rates or valuation ratios have differential exposure to this factor. To shed light on these issues, we analyze firm behavior and endogenous price dynamics in equilibrium, in the spirit of CIR.

In the next section, we develop a simple two-period production economy model building closely on the ideas in [Papanikolaou \(2011\)](#); [Garleanu, Kogan, and Panageas \(2012\)](#); [Kogan and Papanikolaou \(2014\)](#); [Kogan, Papanikolaou, and Stoffman \(2018\)](#).

2.2 An equilibrium model of asset prices with technological change

The goal of our model is to connect the “value effect” in equity returns to technological shocks. The model emphasizes one source of fundamental differences across firms: their heterogeneous ability to benefit from economy-wide technological advances.

Technology and firms

The economy exists for two periods, and there are two vintages of technologies in this model, “old” and “new”. For simplicity, we abstract away from endogenous capital accumulation and assume that output is produced using labor as the sole input. There exists a large set of firms (formally, a continuum of measure one), which collectively have access to both technologies as we discuss below.

Each firm j operates one project using old technology, producing output according to the production function

$$Y_{o,t}^j = (X_t L_{o,t}^j)^{1-\alpha}, \quad t = 0, 1, \quad (2)$$

where $\alpha \in (0, 1)$ and $L_{o,t}^j$ is the corresponding labor input. Labor markets are frictionless, and firms can hire any amount of labor they need, at the equilibrium wage W_t . X_t is the economy-wide labor-augmenting productivity shock. We set $X_0 = 1$.⁵

The new technology is available only in the second period, so in the first period all firms operate only the old technology. We assume that some of the firms operate only the old technology in both periods, while some also have access to the new technology in the second period.⁶

Specifically, we assume that at time $t = 1$ identical new-technology projects arrive in the economy – as many as the old-technology projects. At $t = 1$, each new-technology project j produces output according to

$$Y_{n,1}^j = \xi_1^\alpha (X_1 L_{n,1}^j)^{1-\alpha}, \quad (4)$$

where $\xi_1 \in (0, \bar{\xi})$ is a positive random variable realized at time $t = 1$. ξ_1 represents the vintage-specific technology shock, which affects only the new technology. In models with physical capital, such technological shocks are viewed as embodied in capital of various vintages. A value of $\xi_1 > 1$ implies that the new technology is more productive than the old. A fraction $\omega \in (0, 1)$ of firms own new-technology projects, one project per firm. Thus, collectively these firms introduce a fraction ω of new-technology

⁵While we do not consider firm-specific productivity shocks, these would be straightforward to introduce by assuming the production function of the form

$$Y_{o,t}^j = \left(e_t^j X_t L_{o,t}^j \right)^{1-\alpha}, \quad (3)$$

where e_t^j is the firm-specific productivity shock. If we assume that e_t^j 's are distributed independently of each other and of X_t , the model delivers identical aggregate implications. In the language of APT, the distribution of labor-augmenting productivity shocks then has a single-factor structure.

⁶While we do not take an explicit stand on the economic mechanism behind these differences, one factor that determines whether firms have the opportunity to invest in the new technology is the quality of the firm's management team, as in the model of [Frydman and Papanikolaou \(2018\)](#).

projects into the economy. We assume that this fraction is relatively small:

$$\omega < \frac{1 - \alpha}{1 + \alpha \bar{\xi}}. \quad (5)$$

Last, labor can be flexibly allocated between the old and the new technology.

Investors, workers, and innovators

The economy is populated by two types of agents, investors and workers, who live for two periods. There is a large number of investors with logarithmic preferences over consumption C_0 and C_1 : they maximize

$$\ln C_0 + E_0[\ln C_1]. \quad (6)$$

Investors collectively own all the firms in existence at time $t = 0$. Workers in this economy inelastically supply a fixed amount of labor in each period. Each worker sells labor services to the firms, and does not participate in financial markets. We normalize the aggregate supply of labor to one in each period.

In the second period, a very small number of randomly chosen agents – investors and workers – innovate and create new-technology projects. For simplicity, we assume that these innovators form an infinitesimal fraction of the total population (a zero-measure set). Thus, when an investor formulates her investment and consumption plans at time $t = 0$, she does not expect to innovate in the second period. This assumption simplifies the analysis and does not qualitatively affect the results. Collectively, inventors create a fraction $(1 - \omega)$ of all new-technology projects in the second period – that is, the projects that are not owned directly by the firms (and their shareholders). We assume that innovators monetize their projects by selling them at the fair market value to existing firms. These transactions do not affect firm value.⁷

⁷Alternatively, and without changing any of the model implications, we can assume that innovators

Financial markets

Time-0 investors are marginal in this economy, and thus determine the equilibrium properties of asset returns. We assume that these investors have access to frictionless financial markets, where they can trade all state-contingent claims conditional on the time-1 realizations of technological shocks X_1 and ξ_1 . We denote by π_t the equilibrium stochastic discount factor. Importantly, a key market is missing in this model: time-0 investors cannot trade with the future innovators to share the uncertain benefits of innovation. This assumption can be motivated on theoretical grounds: new ideas are the product of human capital, which is inalienable. This friction is a critical driver of the equilibrium risk premia associated with technological shocks.

Equilibrium

We assume that all agents in this economy behave competitively: firms take prices of state-contingent claims as given and choose their production policies to maximize their market value; investors take prices in financial markets as given and choose their portfolio and consumption policies to maximize their expected utility; workers take wages as given and maximize their labor income. Because the likelihood that a given agent innovates (and receives a fraction of the $1 - \omega$ new-technology projects) is infinitesimally small, all investors behave as if they personally will not innovate at time 1. Accordingly, when we define the combined wealth or consumption of investors, we exclude the wealth and consumption of future innovators.

In equilibrium, the allocation of labor between the old and the new technology

create new firms and directly operate their new-technology projects at time $t = 1$, fully appropriating the economic value added from these projects.

depends on the realization of the embodied shock ξ_1 :

$$L_{o,1}^j = \frac{1}{1 + \xi_1} \quad \text{and} \quad L_{n,1}^j = \frac{\xi_1}{1 + \xi_1}. \quad (7)$$

Labor is paid its marginal product in wages, and because of the functional form of the production technology (Cobb-Douglas), profits of all projects are equal to a fraction α of their output.

In the first period, equilibrium consumption of investors is proportional to the output of the old technology, $C_0 = \alpha$. In the second period, these investors collect profits of all the old-technology projects, as well as a fraction ω of the new-technology projects—which they own by virtue of owning all the firms in the economy at time $t = 0$. As a result, their consumption equals

$$C_1 = \alpha X_1^{1-\alpha} \frac{1 + \omega \xi_1}{(1 + \xi_1)^{1-\alpha}}, \quad (8)$$

and the equilibrium stochastic discount factor (SDF) is given by

$$\frac{\pi_1}{\pi_0} = \frac{C_0}{C_1} = X_1^{\alpha-1} \frac{(1 + \xi_1)^{1-\alpha}}{1 + \omega \xi_1}. \quad (9)$$

The economic content of our assumption (5) is that the owners of the first-generation technology have only a limited claim on the new technology. Being largely tied up in the old-technology projects, their wealth declines in productivity of the new technology. Time-1 equilibrium wealth of investors determines state prices in this economy. Hence, the equilibrium SDF π_1/π_0 is increasing in ξ_1 .

Displacement effect of innovation and asset returns

The two technologies compete at time $t = 1$. In general, this competition may result in the falling prices of output and rising costs of input. In our model, because both technologies produce the same good, the price of output of the old technology at time

$t = 1$ (in units of the consumption good) is fixed at one and not affected by the entry of the new technology. However, the cost of inputs (equilibrium wage) is affected by competition between the two technologies. If the new technology turns out to be highly productive (ξ_1 is high), then the equilibrium wage rises and profits of the projects operating the old technology decline: specifically, the equilibrium return to the old-technology projects,

$$R_1^o = \frac{X_1^{1-\alpha}(1 + \xi_1)^{\alpha-1}}{E_0 [(1 + \omega\xi_1)^{-1}]} - 1, \quad (10)$$

is decreasing in ξ_1 .

Profits of the new-technology projects increase in the higher productivity shock ξ_1 : any increase in equilibrium wages is more than offset by higher productivity of the new technology. In equilibrium, the return on a claim to profits from such projects equals

$$R_1^n = X_1^{1-\alpha} \frac{\xi_1(1 + \xi_1)^{\alpha-1}}{E_0 [\xi_1(1 + \omega\xi_1)^{-1}]} - 1. \quad (11)$$

Unlike the return to old projects, the return to new-technology projects is increasing in ξ_1 .

In summary, we find that the entry of the new technology has a displacive effect on the value of the old-technology projects, and consequently on the stock market wealth of investors. Financial claims that benefit from higher productivity of the new technology are relatively highly valued by investors as a hedge against the displacement effect. These claims earn lower expected rates of return in equilibrium. In particular, based on the above results, the financial claim on the profits of the new-technology projects has a lower expected return than the claim on the profits of the old-technology projects:

$$E_0[R_1^n] < E_0[R_1^o]. \quad (12)$$

By contrast, higher labor-embodied productivity raises profits of both technologies and lowers the equilibrium SDF. The reason is that labor can be flexibly allocated between the two technologies, hence, higher labor productivity has no displacive effect on profits of the old-technology projects. Thus, the labor-augmenting technological shock X_1 is positively compensated with a risk premium in equilibrium, while the embodied shock ξ_1 earns a negative premium.

Stock returns

Firms in our model derive their value from their ownership of old- and new-technology projects. There are two types of firms in our model, which deliver identical cash flows in the first period, but differ by whether they can access the new-technology in the second period. We can think of the claim on the old technology as the firm's assets in place, while the claim on the time-1 new technology represents firm's growth opportunities.

A firm that does not own any new-technology projects is a pure claim on the old technology. In our model, a fraction $(1 - \omega)$ of all firms are of such type. Based on the above analysis, equilibrium return on this firm is given by (10):

$$R_1^v = R_1^o. \tag{13}$$

We refer to such firms as *value firms*: their time-0 valuations are relatively low,

$$P_0^v = \alpha E_0 \left[\frac{1}{1 + \omega \xi_1} \right], \tag{14}$$

and, as we show below, their expected rate of return is relatively high.

A firm that does own a claim on the new-technology projects—a *growth firm*—derives its value both from assets in place and growth opportunities. A fraction ω of

all firms in our model are growth firms. Their time-0 value is relatively high,

$$P_0^g = \alpha E_0 \left[\frac{1 + \xi_1}{1 + \omega \xi_1} \right]. \quad (15)$$

The return on growth firms is a weighted average of returns on the pure old- and new-technology claims,

$$R_1^g = \frac{P_0^v}{P_0^g} R_1^o + \left(1 - \frac{P_0^v}{P_0^g} \right) R_1^n. \quad (16)$$

Thus, according to (12), expected returns on growth firms are relatively low.

We thus obtain a two-factor return model:

$$\ln(1 + R_1^o) = (1 - \alpha) \ln X_1 + (\alpha - 1) \ln(1 + \xi_1) + \text{const}, \quad (17)$$

$$\ln(1 + R_1^n) = (1 - \alpha) \ln X_1 + (\alpha - 1) \ln(1 + \xi_1) + \ln \xi_1 + \text{const}. \quad (18)$$

The two systematic return factors correspond to the two types of productivity shocks: labor-augmenting (X_1) and vintage-specific, or embodied (ξ_1). Linearizing around the expected values of the shocks yields an APT-style model.⁸

All firms in our model have the same loading (APT beta) on the labor-augmenting shock X_1 . Growth firms have positive betas on the embodied shocks ξ_1 , through their ownership of claims on new-technology projects, while value firms have negative betas because of their exposure to displacement risk created by the new technology.

The equilibrium return on the market portfolio is a weighted average of the returns on the two types of firms:

$$R_1^m = \frac{\omega P_0^g}{\omega P_0^g + (1 - \omega) P_0^v} R_1^g + \frac{(1 - \omega) P_0^v}{\omega P_0^g + (1 - \omega) P_0^v} R_1^v. \quad (19)$$

Because any rotation of the two factors yields a valid APT model, we take the market return as the first factor. We form the second factor using a self-financing long-short

⁸There is no idiosyncratic randomness in our model – this is easy to incorporate, see footnote 5.

portfolio of growth and value firms. This portfolio has zero exposure to the labor-augmenting shock, and a positive beta on the embodied shock. It thus serves as a mimicking portfolio for the new technology shock.

Empirical implications

We now relate the properties of our stylized theoretical model to the empirical patterns described in Section 2.1. We first provide evidence that empirical proxies for technological changes (the new-technology shocks in our model) correlate with aggregate quantities and prices in a manner that is consistent with the theory. To do so, we use the empirical measure of the value of new technologies developed in [Kogan, Papanikolaou, Seru, and Stoffman \(2014\)](#) and [Kogan et al. \(2018\)](#). This measure combines information in patents and stock market data.

Table 4 shows that our innovation measure correlates with aggregate investment growth, the market portfolio, the value factor, and the PC1 pricing factor discussed in Section 2. These correlation patterns support the model predictions that the value of the aggregate stock market is negatively impacted by innovation, and that innovation shocks have a stronger negative effect on prices of value stocks compared to growth stocks. Moreover, technological innovation shocks generate comovement among stocks in characteristic-sorted portfolios, as evidenced by the positive correlation between innovation shocks and the PC1 return factor.

Next, we consider how innovation shocks correlate with empirical measures of income and wealth inequality. Imperfect sharing of the benefits of technological innovation is an essential element of our model. Because investors in the stock market can only appropriate a fraction ω of the benefits of the new technology, their wealth (and therefore consumption) is negatively impacted by innovation shocks— even though these shocks lead to higher economic growth in the aggregate. Observationally,

Table 4: Innovation, aggregate quantities and asset returns

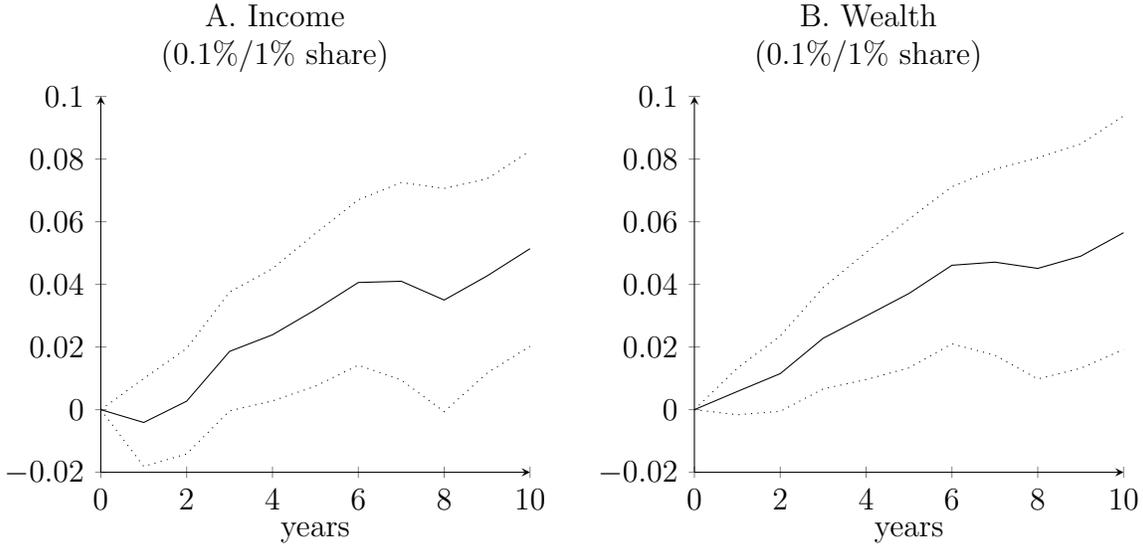
Value of new technologies	Correlation		
	$(t \rightarrow t + 1)$	$(t \rightarrow t + 2)$	$(t \rightarrow t + 3)$
A. Investment growth	0.30 (0.12)	0.27 (0.10)	0.30 (0.12)
B. Market portfolio	-0.50 (0.11)	-0.30 (0.13)	-0.09 (0.13)
C. Value factor	-0.34 (0.09)	-0.44 (0.09)	-0.52 (0.12)
D. PC1	0.27 (0.18)	0.46 (0.12)	0.50 (0.09)

This table summarizes correlations between differences in the rate of innovation and returns to the market portfolio, the value factor (the return spread between the top and bottom decile portfolios of stocks sorted on book-to-market), aggregate investment growth (NIPA), and the PC1 pricing factor in Table 2. For portfolio returns, we compute the cumulative portfolio return between t and $t + h$. For all other variables, including innovation, we use the log difference between t and $t + h$. We compute correlations at horizons of 1 to 3 years. Numbers in parentheses are standard errors, computed using the Newey-West method with a maximum lag length equal to three plus the number of overlapping observations. Data period is 1933 to 2008. See [Kogan and Papanikolaou \(2013\)](#) and [Kogan et al. \(2018\)](#) for further details.

imperfect risk sharing may manifest as rising inequality: as the benefits of the new technology accrue to a small fraction of agents, disparities in income and wealth increase following higher returns to new technologies.

We use two empirical measures of inequality among households. The first of these measures is the ratio of the 0.1% top income share to the top 1%, based on the data of [Piketty and Saez \(2003\)](#). This “fractal” measure of inequality captures the share of the top 1% that accrues to the top 0.1%, and displays similar behavior as the top income shares ([Jones and Kim, 2014](#); [Gabaix, Lasry, Lions, and Moll, 2016](#)). The second measure is equal to the ratio of the top 0.1% wealth share to the top 1%, and uses the data in [Saez and Zucman \(2016\)](#). These two measures of inequality have the advantage of being available over long periods of time, which is important for identifying long-horizon relations. For each of the two inequality measures, we

Figure 1: Technology shocks and inequality



This figure plots the estimated relation between top income and wealth inequality and our measure of technological change based on [Kogan et al. \(2018\)](#). For income, we use the series of top income shares of [Piketty and Saez \(2003\)](#) that include capital gains. For wealth, we use the series on top wealth shares by [Saez and Zucman \(2016\)](#). We plot empirical point estimates along with the 90% confidence intervals. See [Kogan et al. \(2018\)](#) for additional details.

estimate the following regression specification:

$$\log INEQ_T - \log INEQ_t = a_0 + b_T (\hat{\omega}_T - \hat{\omega}_t) + c INEQ_t + d \hat{\omega}_t + u_t, \quad (20)$$

where $\hat{\omega}$ is the empirical measure of innovation.⁹ Results in [Figure 1](#) show that higher rates of innovation in the economy are followed by rising inequality, measured by either income or wealth inequality. This is consistent with the broader patterns of wealth concentration in the U.S. being related to business creation: for instance, [Cagetti and DeNardi \(2006\)](#) find that 81 percent of individuals the top one percent of the U.S. wealth distribution declare to be either self-employed or business owners.

We finally turn to the relation among firm characteristics, risk premia, and return comovement. In the model, firm characteristics contain information about their return

⁹We define $\hat{\omega}$ as the log value of new patents in the data over the total stock market capitalization.

betas on the new-technology shock. Specifically, firms' P/E ratios identify growth and value firm types, with growth firms loading higher on new technology. This result gives rise to two important patterns in returns. First, growth firms earn lower returns on average relative to value firms. Second, returns on high-growth firms comove with each other, after controlling for their market risk exposures, because of their common loading on the new-technology productivity shock.

Table 5: Characteristics, comovement, and risk premia in simulated data

	10 minus 1 decile portfolios				
	Tobin's q	I/K	P/E	MBETA	IVOL
$E(R_p)$ (%)	-5.27 (-5.71)	-4.81 (-7.85)	-7.21 (-3.93)	-5.55 (-4.06)	-5.24 (-5.38)
σ (%)	6.89	4.41	13.33	9.84	7.03
β^{mkt}	0.31 (7.71)	0.22 (7.10)	0.74 (9.53)	0.54 (9.21)	0.36 (6.99)
α (%)	-7.13 (-10.83)	-6.23 (-13.88)	-11.94 (-10.47)	-8.98 (-10.14)	-7.56 (-10.68)
R^2 (%)	53.34	53.01	64.23	63.49	54.79

This table replicates the analysis in Table 1 in simulated data. We compute the mean and volatility of portfolio excess returns, and the results of CAPM regressions. We compute median point estimates of the parameters and the t -statistics across 1,000 simulations, each with length of 50 years. See [Kogan and Papanikolaou \(2013\)](#) for further details.

In our empirical analysis, we consider a broader set of firm characteristics. In addition to valuation ratios, such as price-to-earnings and price-to-book ratios, we sort firms on their past investment, market betas, and idiosyncratic volatility. As we show in [Kogan and Papanikolaou \(2014\)](#), using a dynamic structural model, such firm characteristics contain information about the fraction of firm value arising from firm's future growth opportunities, as opposed to assets in place. For instance, firms with more growth opportunities are likely to invest more, have higher valuation ratios, and

be more exposed to growth-related shocks. Our structural model replicates closely the empirical patterns in Section 2.1. We reproduce the model-based results in Table 5, where we apply empirical procedures to the data simulated from the model. The model generates differences in risk premia across the decile portfolios that are comparable to the data; the differences in average return between the top and bottom decile portfolios range from -4.8% for the I/K sort to -7.2% for the P/E sort. Further, the model reproduces the empirical failure of the CAPM, generating substantial alphas (-6.2% to -11.9%).

3 Conclusion

The empirical facts and the model described in the previous section bring together ideas from both the APT and CIR. Specifically, they show how a theory that explicitly models firm production decisions as in CIR can help us understand the economic content behind APT-style return factors. Given the rising popularity of “smart beta”, that is, investment strategies that are designed to provide investors with targeted exposures to a select number of return factors—with value being the most prominent example—understanding the economic risks that lie behind these factors is of first-order importance.

References

- Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2006). The cross-section of volatility and expected returns. *Journal of Finance* 61(1), 259–299.
- Black, F., M. C. Jensen, and M. S. Scholes (1972). *Studies in the Theory of Capital Markets*, Chapter The Capital Asset Pricing Model: Some Empirical Tests. Praeger Publishers Inc.
- Boldrin, M., L. J. Christiano, and J. D. M. Fisher (2001). Habit persistence, asset returns, and the business cycle. *American Economic Review* 91(1), 149–166.
- Brock, W. A. (1982). Asset Prices in a Production Economy. In *The Economics of Information and Uncertainty*, NBER Chapters, pp. 1–46. National Bureau of Economic Research, Inc.
- Cagetti, M. and M. DeNardi (2006). Entrepreneurship, frictions, and wealth. *Journal of Political Economy* 114(5), 835–870.
- Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *Journal of Finance* 47(2), 427–65.
- Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33(1), 3–56.
- Fama, E. F. and K. R. French (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance* 51(1), 55–84.
- Frazzini, A. and L. Pedersen (2014). Betting against beta. *Journal of Financial Economics* 111(1), 1–25.
- Frydman, C. and D. Papanikolaou (2018). In search of ideas: Technological innovation and executive pay inequality. *Journal of Financial Economics* forthcoming.

- Gabaix, X., J.-M. Lasry, P.-L. Lions, and B. Moll (2016). The dynamics of inequality. *Econometrica* 84(6), 2071–2111.
- Garleanu, N., L. Kogan, and S. Panageas (2012). Displacement risk and asset returns. *Journal of Financial Economics* 105(3), 491–510.
- Gibbons, M. R., S. A. Ross, and J. Shanken (1989). A test of the efficiency of a given portfolio. *Econometrica* 57(5), 1121–1152.
- Jermann, U. J. (1998). Asset pricing in production economies. *Journal of Monetary Economics* 41(2), 257–275.
- Jones, C. I. and J. Kim (2014, October). A schumpeterian model of top income inequality. Working Paper 20637, National Bureau of Economic Research.
- Kaltenbrunner, G. and L. A. Lochstoer (2010). Long-run risk through consumption smoothing. *Review of Financial Studies* 23(8), 3190–3224.
- Kogan, L. and D. Papanikolaou (2013). Firm characteristics and stock returns: The role of investment-specific shocks. *The Review of Financial Studies* 26(11), 2718–2759.
- Kogan, L. and D. Papanikolaou (2014). Growth opportunities, technology shocks, and asset prices. *Journal of Finance* 69(2).
- Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2014). Growth opportunities, technology shocks, and asset prices. *Journal of Finance* 69(2), 675–718.
- Kogan, L., D. Papanikolaou, and N. Stoffman (2018). Winners and losers: Creative destruction and the stock market. *Journal of Political Economy* forthcoming.
- Lakonishok, J., A. Shleifer, and R. W. Vishny (1994). Contrarian investment, extrapolation, and risk. *Journal of Finance* 49(5), 1541–78.

- Papanikolaou, D. (2011). Investment shocks and asset prices. *Journal of Political Economy* 119(4), pp. 639–685.
- Piketty, T. and E. Saez (2003). Income inequality in the united states, 1913—1998. *The Quarterly Journal of Economics* 118(1), 1–41.
- Prescott, E. and R. Mehra (1980). Recursive competitive equilibrium: The case of homogeneous households. *Econometrica* 48(6), 1365–79.
- Rosenberg, B., K. Reid, and R. Lanstein (1985). Persuasive evidence of market inefficiency. *Journal of Portfolio Management* 9, 18–28.
- Saez, E. and G. Zucman (2016). Wealth inequality in the united states since 1913: Evidence from capitalized income tax data. *The Quarterly Journal of Economics*, 519–578.
- Tallarini, T. D. (2000). Risk-sensitive real business cycles. *Journal of Monetary Economics* 45(3), 507 – 532.
- Titman, S., K. C. J. Wei, and F. Xie (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis* 39(4), 677–700.