

UCLA

UCLA Electronic Theses and Dissertations

Title

Essays on Asset Pricing

Permalink

<https://escholarship.org/uc/item/2qz6c3hc>

Author

Cuevas Rodriguez, Gabriel Ignacio

Publication Date

2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Essays on Asset Pricing

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Management

by

Gabriel Ignacio Cuevas Rodriguez

2023

© Copyright by
Gabriel Ignacio Cuevas Rodriguez
2023

ABSTRACT OF THE DISSERTATION

Essays on Asset Pricing

by

Gabriel Ignacio Cuevas Rodriguez

Doctor of Philosophy in Management

University of California, Los Angeles, 2023

Professor Bernard Herskovic, Co-Chair

Professor Stavros Panageas, Co-Chair

In Chapter 1, I analyze firms' misallocation through the output distortions channel, using a production-based asset pricing model as a framework. In the model, α measures the firm's ability to choose technologies to adapt to exogenous shocks. I find in the cross-section of the test portfolios the estimated curvature parameter α is more than two times the original value obtained in Belo (2010). This implies misallocations reduce the firm's ability to respond to the different states of nature. I calibrate and solve the model in the special case of a single representative firm. I find that the impact of misallocation on firm value, production, capital, investment, and investment return is larger when firms' ability to adapt to exogenous shocks is reduced. This indicates that firms may be less agile to adapt across states of nature and provides more evidence of the detrimental effect of misallocations.

In Chapter 2 (with Denis Moganov and Danyu Zhang), we document several facts

about equity analysts' earnings expectations: (1) consensus earnings expectations underreact to news unconditionally, (2) the degree of underreaction declines during high-volatility periods, and (3) the degree of underreaction declines over our sample. To account for these findings, we develop a simple model featuring time-varying inattention. We show that our model is able to account for the unconditional profitability of momentum, momentum crashes, and the diminishing profitability of momentum over our sample. We propose a trading strategy that mixes short-run and long-run momentum signals and show that the mixed momentum strategy outperforms the conventional momentum strategies. Finally, we use a machine learning algorithm to estimate the predictable component of earnings surprises and construct a portfolio that is long (short) on stocks with excessively pessimistic (optimistic) earnings expectations. The resultant trading strategy generates an annualized Sharpe ratio of about 1.16 and its returns are not explained by popular factor models.

The dissertation of Gabriel Ignacio Cuevas Rodriguez is approved.

Valentin P. Haddad

Mikhail Chernov

Stavros Panageas, Committee Co-Chair

Bernard Herskovic, Committee Co-Chair

University of California, Los Angeles

2023

To my mother Rudy, Maria and Emma

TABLE OF CONTENTS

1	The Effect of Misallocation on Asset Prices	1
1.1	Introduction	1
1.1.1	Related literature	3
1.2	A two-period production-based model	5
1.2.1	Productivity	5
1.2.2	Firm's distortions	8
1.2.3	Producer's first-order conditions	9
1.3	A multi-period production-based model	10
1.3.1	Firms	11
1.3.2	The producer's maximization problem	13
1.3.3	First-order conditions	13
1.3.4	Identification	15
1.3.5	Asset Pricing Implications	17
1.4	Quantitative Analysis	18
1.4.1	Data	18
1.4.2	Estimation	20
1.4.3	Results	22
1.4.4	Results: Additional Test Assets	26
1.5	Simulation Analysis	29

1.5.1	Explicit Representation of a Single Firm	29
1.5.2	Numerical Approach	32
1.5.3	Scenario Analysis	32
1.5.4	Comparative Statics of Contingent Distortions	38
1.6	Conclusion	43
Appendices		45
1.A	Proofs and Derivation	45
1.A.1	Producer's Maximization Problem	45
2	Earnings Expectations and Asset Prices	51
2.1	Introduction	51
2.2	Data	56
2.2.1	Analysts' forecasts	56
2.2.2	Stock and trading strategy returns	57
2.3	Forecast Error Predictability	58
2.4	Model	62
2.4.1	Portfolio choice problem	64
2.4.2	Information acquisition problem	65
2.4.3	Equilibrium	66
2.4.4	Model implications for forecast error predictability	68
2.5	Asset Pricing Implications	70

2.5.1	Momentum	71
2.5.2	Momentum and volatility	72
2.5.3	Attenuation of momentum	78
2.5.4	Return Predictability	80
2.6	Conclusion	89
Appendices		103
2.A	Derivations	103
2.A.1	Information Acquisition Problem	103
2.A.2	Information Rigidity Coefficient	104
2.A.3	Comparative Statics	105
2.A.4	Momentum	106
2.B	Robustness: Jensen-Kelly-Pedersen Momentum Factor	107
2.B.1	Momentum and Volatility	107
2.B.2	Mixed Momentum	108
2.C	XGBoost	108

LIST OF FIGURES

1.1	Production Possibility Example	6
1.2	Pricing Errors	25
1.3	Firm's Value	35
1.4	Output	36
1.5	Capital	37
1.6	Investment	39
1.7	Dividend	40
1.8	Expected Excess Investment Return	41
1.9	Comparative Statistics Misallocation vs Standard Model	43
2.1	Information Rigidity Over Time	96
2.2	Model Representation	97
2.3	Average Price Path	98
2.4	Average Price Paths, High and Low Volatility States	99
2.5	Mixed Momentum Strategy	100
2.6	Cost of Information and Momentum	101
2.7	Decision Tree Example	102
2.8	Mixed Momentum Strategy	111

LIST OF TABLES

1.1	Summary statistics.	19
1.2	GMM estimation of the production-based model.	23
1.3	Additional Test Portfolio Estimation	28
1.4	Calibrated and Estimated Parameter Values	33
1.5	Misallocation Scenario Values	33
2.1	Forecast Error Predictability Regressions	91
2.2	Momentum and Volatility	92
2.3	XGBoost Hyper-parameters	93
2.4	Mixed Momentum Spanning Regressions	93
2.5	Portfolios Sorted on Expectation Errors	94
2.6	PMO Spanning Regressions	95
2.7	Momentum and Volatility	110

ACKNOWLEDGMENTS

I will start by thanking those whose feedback directly influenced the dissertation. First is my doctoral committee, including my chairs Bernard Herskovic and Stavros Panageas, along with Mikhail Chernov and Valentin Haddad. Other finance faculty who have positively impacted my time at UCLA include Andrea Eisfeldt, Mark Garmaise, Barney Hartman-Glaser, Lars Lochstoer, Francis Longstaff, Tyler Muir, and Ivo Welch,

I am also indebted to Eduardo Schwartz, for his support and for being an inspiration to pursue this degree.

I have also had the pleasure of working with a talented group of peers that includes Salil Gadgil, Chady Gemayel, Ljubica Georgievska, Paul Huebner, Ed Kim, Wenyu Meng, Denis Mokanov, Paul Momtaz, James O'Neill, Zach Sauers, Yu Shi, Clinton Tepper, Daniel Yavorsky, Danyu Zhang, and Geoffery Zheng. Special thanks to Matias Giacobasso for becoming a lifelong friend.

Thank you to UCLA Anderson current and former staff, including Veronica Avila, Maralmaa Chinbat, Kristin Christian, Craig Jessen, Monica Pena, and DaJanay Ray for their help.

Lastly, I thank the UCLA Graduate Division, the UCLA Anderson Finance area, and the UCLA Fink Center for Finance for generous funding and financial support during my time completing this research.

VITA

- 2009 B.S. in Electrical Engineering, Universidad de Chile.
- 2011 B.S. in Industrial Engineering, Universidad de Chile.
- 2016 Master of Financial Engineering, UCLA Anderson School of Management.
- 2016–2023 Teaching Assistant, UCLA Anderson School of Management. Taught Stochastic Calculus and Derivatives, Corporate Finance, Empirical Methods in Finance, Credit Markets, Financial Market Risk and Regulation, Foundations of Finance, Real Estate Trends: Data Analytics & PropTech, Venture Capital and Private Equity, Introduction to Finance of Blockchain and Cryptocurrency, Options Markets, and Derivative Markets.
- 2017–2021 Fink Center Scholar.
- 2017–2021 Anderson Fellow.
- 2017–2023 Ph.D. Candidate in Management (Finance), UCLA Anderson School of Management.

CHAPTER 1

The Effect of Misallocation on Asset Prices

1.1 Introduction

The study of asset pricing invariably involves analyzing the properties of the stochastic discount factor (SDF) as dictated by $P_t = \mathbb{E}_t[M_{t+1}X_{t+1}]$, where X_{t+1} consists of the future payoffs for a given asset or portfolio and M_{t+1} is a vector Arrow-Debreu prices for all states of nature. As discussed at length in Cochrane (1996) and Cochrane (1991), the investment returns of firms correspond to the returns on financial assets. Firms make investment decisions to maximize profits. The firm's beneficial owners realize these returns and price the cashflows in expectation according to the market's SDF.¹

Departing from the traditional consumption-based asset pricing literature, Cochrane (1991) proposes a partial equilibrium model that omits consumer preferences but still specifies the SDF from the producers' first-order conditions. As pointed out by Cochrane, the empirical advantage of the production-based models, contrary to the consumption-based literature, lies in the relatively large movements of output or in-

1. The analysis of the SDF in the context of production forms the basis of production-based asset pricing. See Chapter 7 of Campbell (2017) for an excellent survey of the existing production-based asset pricing literature.

vestment compared to consumption. For instance, the equity premium puzzle results in large part because of consumption smoothness that confounds the reconciliation of the theory with data.

As discussed by Cochrane, when markets are complete, the return on investment must be equal to any portfolio return that mimics the investment return in all states of nature. More importantly, this SDF has to be unique. However, firms are distant from being unique, and idiosyncracies in the form of misallocations can explain a large amount of differences in output. For instance, Restuccia and Rogerson (2008) find that misallocation can cause quantitatively large output and productivity losses, on the order of 30 to 50%. Hsieh and Klenow (2009) find that differences in total factor productivity (TFP) for China and India compared to the U.S., can account for more than two times the TFP. This paper explores if misallocations are priced in the cross-section.

Can distortions in firms' profits influence not only firm investment decisions and production allocations across states of nature, but also asset prices in the cross-section? The model presented in Belo (2010) provides a parsimonious vehicle to consider such effects. Belo directly specifies the marginal rate of transformation (MRT) implied by the firm's production possibility frontier.² He shows that the MRT forms a valid SDF that prices stock returns in the cross-section. Because his specification involves choosing investments across states of nature, as opposed to working at the product or the project level, this model forms an ideal vehicle for subsequent analysis studying the impact of distortions.

2. Marginal rates of transformation are equivalent from the production perspective to the marginal rates of substitutions inferred from consumers' first-order conditions.

While the Belo (2010) model provides a framework, the original model contains no misallocation, and the application of state-contingent distortions in a production framework is novel to the best of my knowledge. Belo (2010) harnesses the production model to empirically estimate asset pricing parameters. I take a similar approach and estimate the parameters of the model using standard Fama French portfolios. In Belo’s model, α measures the firm’s ability to choose technologies to adapt to exogenous shocks. I find a cross-sectional average α of 3.79, indicating the firm’s adaptability to switch across states is lower compared to Belo’s original model ($\alpha = 1.02$). Another measure commonly used in the literature is the mean absolute pricing errors (MAPE). I find the model with misallocation consistently reduces the pricing errors, with a cross-section average reduction of 7 bps. Finally, I calibrate the model and present the results of a simulation exercise, where I actively solve the model in the special case of a single representative firm. The model’s solution generates insights into the effect of misallocation on relative production allocations, investment, and value. The results indicate that misallocation has a detrimental impact on firms. Firms have a lower ability to adapt to exogenous shocks and as a consequence, their value is also reduced.

1.1.1 Related literature

This paper extends the literature on production-based asset pricing. The goal of this approach is to understand how misallocation can impact the firm’s decisions and consequently asset prices. This paper relates to the literature that focuses on the production side of the economy as a mirror to consumption-based models in the asset pricing literature. The first efforts were made by Cochrane (1988), focusing

on a smooth production possibility frontier across states of nature. Similarly to the consumption-based approach, in which consumers save to increase their consumption in states of nature where the marginal rate of substitution is high, a firm can take actions to shift their output from one state to another. Cochrane (1988) has recently received more attention in a direct extension by Belo (2010) and Cochrane (2021).

The work most closely related to mine is Belo (2010). Belo builds on Cochrane (1993) and proposes a production-based asset pricing model, based on the ability of the producers to shift output across states. His approach departs from the usual aggregate production function and requires that firms have a smooth choice over the state-contingent pattern of their output. A smooth production possibility frontier makes it possible to obtain the MRT, which must be equal to the SDF in each state. The model has an unobservable parameter, the natural productivity. Belo solves this identification problem by assuming it follows a factor structure. I base my model on Belo's study, but I differ by considering that firms have capital misallocation represented as output distortions. Cochrane (2021) presents several variations of the production-based asset-pricing model, but the goal of that paper is to establish a research agenda more than to propose answers to the production-based literature. The use of a smooth production possibility frontier has also proven to be flexible enough to relate asset prices and the firm's research and development (R&D). Guo, Zhang, and Zhang (2022) explore the possibility that the technology component can be explained using R&D. In particular, they develop a technique to identify natural productivity shocks through technological advancements.

This paper is also related to the literature on misallocation. Restuccia and Rogerson (2008) explore a stationary equilibrium, while assuming a heterogeneous set of firms, with constant TFP and misallocations in the form of tax rates. They find

that misallocation can cause quantitatively large output and productivity losses, on the order of 30 to 50%. Distortions may or may not be correlated with firm size, generally finding larger effects if correlated with size. I depart from Restuccia and Rogerson (2008), as I assume that distortions are not constant. Hsieh and Klenow (2009) also find that distortions affect TFP. Their empirical analysis is at a country level, but their findings imply that large differences in TFP can be explained by inefficient uses of technologies (licensing regulations, size-dependent policies, state-owned enterprises). I apply a similar methodology to estimate distortions, but since my model abstracts from labor, I only consider capital distortions. Finally, Bloom et al. (2018) study uncertainty in the business cycle. They find that uncertainty is countercyclical, and one of the channels is misallocation. Following their findings, I consider state-dependent distortions.

1.2 A two-period production-based model

In this section, I consider a simple model to make the dynamics of output, the SDF, and misallocation transparent. There are two periods and a representative firm that has the possibility to invest at time 0.

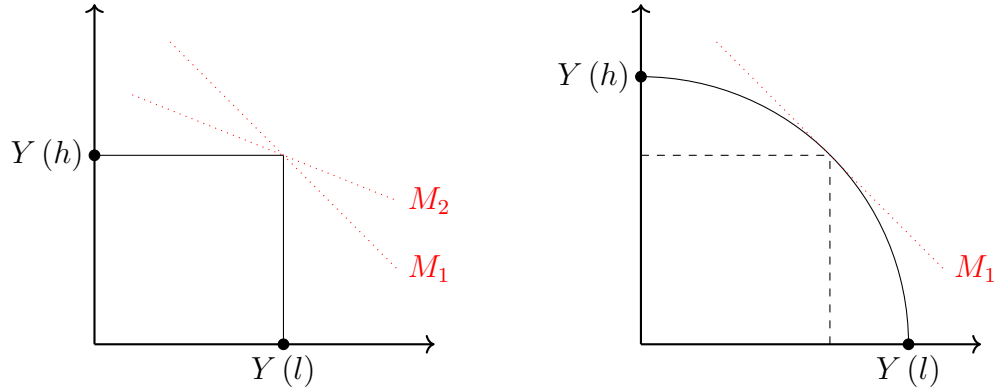
1.2.1 Productivity

In an uncertain environment, the output is

$$Y(s_1) = \varepsilon(s_1) F(K_0) \tag{1.1}$$

where $\varepsilon(s_1)$ is random productivity in state of nature s_1 at time 1, and $F(K_0)$ is the production function on capital. The firm invests K_0 at time 0 before the shock ε is realized. Under standard aggregate firm's representation, the firm has the ability to adjust its output over time, but not transform it across states of nature.³ As a result, there is a kink in the production possibilities frontier, and the MRT across states is not well defined. Figure 1.1 (left) displays a simple representation of the production possibility frontier for a two-state $s = h, l$ economy. We can observe that the MRT is not well defined and two possible MRT M_1 and M_2 are displayed at the same production level.

Figure 1.1
Production Possibility Example



Production possibility set for a two-state economy: (i) left: standard; (ii) right: smooth production possibility set.

Following Cochrane (1993) and Cochrane (2021) and Belo (2010), my goal is to use a production technology that in addition to the usual ability to transform

3. The resulting production possibilities frontier is Leontief. Cochrane (1993) and Cochrane (2021), and Belo (2010) discuss this in detail.

goods over time, allows firms to adapt across states of nature. This representation is determined by a standard technology as in Eq. (1.1), but the firm has the ability to choose a state-contingent productivity level $\varepsilon(s_1)$, subject to a constraint set of the form of a CES aggregator:⁴

$$\left(\mathbb{E} \left[\left(\frac{\varepsilon(s_1)}{\theta(s_1)} \right)^\alpha \right] \right)^{\frac{1}{\alpha}} \leq 1, \quad (1.2)$$

where $\alpha > 1$ is a curvature parameter, and θ are a set of weights. The restriction on α warrants the concavity of the production possibility frontier.⁵ θ is a state-contingent technological parameter, that can be thought of as underlying random productivity. θ allows the firm to obtain higher productivity in some states, by accepting lower productivity in other states. Figure 1.1 (right) displays a simple representation of a smooth production possibility frontier. In this case, the MRT M_1 is well-defined and unique.

It is convenient to understand that θ is not directly observable. Assuming that there is a finite number of states $s_1 = 1, 2 \dots S$, Eq. (1.2) becomes⁶

$$\left(\sum_{s_1=1}^S \rho(s_1) \left(\frac{\varepsilon(s_1)}{\theta(s_1)} \right)^\alpha \right)^{\frac{1}{\alpha}} \leq 1. \quad (1.3)$$

where $\rho(s_1)$ represents the probability of each state $s_1 = 1, \dots, S$. The parameters

4. Belo (2007)'s appendix A.II describes in detail the steps to arrive to Eq. (1.2) from the firm's output.

5. When α approaches 1, it is easier for the firm to transform output across states of nature. $\alpha \rightarrow \infty$ recovers the Leontief representation of the production possibility frontier.

6. The finite state example is for simplicity but not required as discussed by Cochrane (2021)

θ and ρ are not separately identified, so a change in one can be compensated by the other. It is important to notice that from the firm's perspective, it is easier to produce in states of nature where θ is high, and more difficult when θ is low. The hypothesis that firms have smooth opportunity sets, and thus have some control over nature, is discussed in detail in Cochrane (2021), and Belo (2010).

1.2.2 Firm's distortions

Following Belo (2010), I abstract from labor and adjustment costs to keep the model simple. This assumption allows me to introduce a unique distortion that changes the marginal product of capital in this economy. Since capital is the only factor of production, there is no need to add additional distortions that may affect the other factors of production. Similar to Hsieh and Klenow (2009), I denote distortions on capital as an output distortion τ .⁷ For instance, as discussed by Hsieh and Klenow (2009), τ would be high for firms that face government restrictions, and low for firms that benefit from subsidies. The firm's profits are given by

$$\Pi(s_1) = Y(s_1)(1 - \tau(s_1)). \quad (1.4)$$

Therefore, at time 0, the firm's planning problem is

$$\max_{K_0, \varepsilon(s_1)} \mathbb{E}[M(s_1)\Pi(s_1)] - K_0 \quad \text{s.t.} \quad \left(\mathbb{E} \left[\left(\frac{\varepsilon(s_1)}{\theta(s_1)} \right)^\alpha \right] \right)^{\frac{1}{\alpha}} \leq 1. \quad (1.5)$$

7. Hsieh and Klenow (2009)'s model has two factors of production: capital and labor. Thus, they introduce two distortions: one for capital and labor together by the same proportion, as an output distortion; the second distortion is the capital relative to labor. This representation allows them to identify separately the impact on each factor of production.

where M is the SDF.

1.2.3 Producer's first-order conditions

Introducing a Lagrange multiplier μ on the productivity-choice constraint of (1.5), the first-order conditions are

$$\frac{\partial}{\partial K_0} : 1 = \mathbb{E}[M(s_1) \varepsilon(s_1) F'(K_0) (1 - \tau(s_1))] \quad (1.6)$$

$$\frac{\partial}{\partial \varepsilon(s_1)} : M(s_1) F(K_0) (1 - \tau(s_1)) = \mu \frac{\varepsilon(s_1)^{\alpha-1}}{\theta(s_1)^\alpha} \quad (1.7)$$

where I use that firms always maximize production, so the constraint in Eq. (1.5) is binding. The firm should invest until the physical investment return is correctly priced. Thus, an alternative representation of Eq. (1.6) is⁸

$$1 = \mathbb{E}[MR^I] \quad (1.8)$$

with $R^I = \varepsilon F'(K_0) (1 - \tau)$, denoting the investment return. Eq. (1.8) implies the familiar result from Cochrane (1991): the law of one price implies production returns consistent with the returns of the firm's equity holders. While the model itself taken literally is untestable due to misspecification and measurement error, Cochrane and others develop interesting and novel techniques to test the theory in principle. Such testing is tangential to the subsequent analysis, which examines the partial equilibrium implications of production misallocations.

8. For simplicity, I omit the state-contingent-dependency s_1 for the state-contingent variables.

Now, solving for the SDF in the productivity choice first-order condition,

$$M = \frac{\varepsilon^{\alpha-1}}{(1-\tau) F' (K_0) \theta^\alpha} = \frac{1}{R^I} \left(\frac{\varepsilon}{\theta} \right)^\alpha \quad (1.9)$$

Eq. (1.9) says that firms will have high discount factors when productivity is high and when output distortions are larger. Representing Eq. (1.9) in terms of ε

$$\varepsilon = M^{\frac{1}{\alpha-1}} (1-\tau)^{\frac{1}{\alpha-1}} F' (K_0)^{\frac{1}{\alpha-1}} \theta^{\frac{\alpha}{\alpha-1}}, \quad (1.10)$$

implies that firms' productivity level is determined by technological constraints and capital distortions. Given that $\alpha > 1$, the firm chooses a high productivity level when m is high; in states when it is easier to produce, high θ states; and when distortions are reduced, low τ states. This is in line with the findings of Restuccia and Rogerson (2008), where the larger the idiosyncratic distortions, the lower the relative output.⁹

The upshot is an empirically tractable production factor and SDF. The SDF in this case represents the MRT, and should effectively price the returns from future production.

In the next section, I present a multi-period generalization of this model.

1.3 A multi-period production-based model

This section presents a production-based asset pricing model in partial equilibrium. Consider the model presented in Belo (2010) with one significant change: firms

9. Hsieh and Klenow (2009) obtain similar findings with distortions associated with the two productivity factors.

face state-dependent output distortions. The firm's smooth production possibility frontier allows for well-defined MRT. Firms' first-order conditions equate the SDF to the MRT in each state of nature. Finally, following Belo (2010), I propose a methodology to identify the SDF in the data, I validate the SDF in the cross-section using standard Fama French test portfolios.

1.3.1 Firms

Competitive firms take prices as given and produce a single final good using capital. Output, Y_{jt+1} , of firm j is produced by a standard technology with capital, K_{jt+1} , as the only factor input:

$$Y_{jt+1} = \varepsilon_{jt+1} F^j(K_{jt+1}), \quad (1.11)$$

where $F^j(\cdot)$ is an increasing and concave function of the inputs K_{jt+1} .

In addition, following Cochrane (1993) and Cochrane (2021) and Belo (2010), each firm picks a state-dependent production level ε_{jt+1} for all states of nature. The producer chooses their production level subject to the efficacy of their production process in each state. Using a CES aggregator, the constraint takes the form:

$$\left(\mathbb{E}_t \left[\left(\frac{\varepsilon_{jt+1}}{\theta_{jt+1}} \right)^\alpha \right] \right)^{\frac{1}{\alpha}} \leq 1 \quad (1.12)$$

where θ_{jt+1} is the technology vector for firm j across all states. Note that θ_{jt+1} and ε_{jt+1} are random, contingent on the state realization. The parameter α determines the flexibility of technology. When $\alpha \rightarrow \infty$, firms cannot transform output, recovering the standard case of a Leontief across the states of nature. In addition, $\alpha > 1$ is

a necessary restriction that guarantees a concave frontier across states of nature.

At time t , firm j makes an investment, I_{jt} , to the capital stock, K_{jt+1} , next period. The capital accumulation equation is given by:

$$K_{jt+1} = I_{jt} + (1 - \delta) K_{jt} \quad (1.13)$$

where δ is the depreciation rate of the producer's capital stock.

At time t , the firm pays out dividends given by

$$D_{jt} = P_{jt} Y_{jt} (1 - \tau_{jt}) - I_{jt}, \quad (1.14)$$

where P_{jt} is the price of the firm's goods at time t , and τ_{jt} represent the state-contingent output distortions. The assumption that output distortions are state-contingent is based on Bloom et al. (2018) findings that uncertainty rises during recessions. Bloom et al. (2018) propose a general equilibrium model, where the volatility of the uncertainty in output follows a two-state Markov chain.¹⁰ They find that the volatility of the two processes increases during recessions. In their model, one of the channels for the idiosyncratic process is misallocation, which rises during recessions. Additionally, as discussed in Section 1.2, each firm has a unique output distortion given by τ_{jt} . The output distortion is in line with Hsieh and Klenow (2009) model, but since I abstract from a labor productivity factor, I assume that the distortion to capital affects directly output.¹¹

10. Bloom et al. (2018) define uncertainty as two processes: one is a macroeconomic process, and the second is an idiosyncratic process.

11. Belo (2010) shows that labor does not have implications in the MRT.

1.3.2 The producer's maximization problem

The firms take the SDF, M_{t+1} , measured in units of a numeraire good, and the relative price of its output $P_{jt} = p_{jt}/p_{it}$, as given.¹² Markets are complete and thus the SDF is unique.¹³ The firm therefore solves

$$V(\mathbf{X}_{jt}) = \max_{I_{jt}, \varepsilon_{jt+1}} \{D_{jt} + \mathbb{E}_t[M_{t+1}V(\mathbf{X}_{jt+1})]\} \quad (1.15)$$

where $\mathbf{X}_{jt} \equiv (K_{jt}, \varepsilon_{jt}, P_{jt}, \mathbf{Z}_{jt})$ and \mathbf{Z}_{jt} contains all forecasting variables.

The firm is subject to the following constraints:

$$D_{jt} = P_{jt}Y_{jt}(1 - \tau_{jt}) - I_{jt} \quad (1.16)$$

$$K_{jt+1} = I_{jt} + (1 - \delta)K_{jt} \quad (1.17)$$

$$1 \geq \left(\mathbb{E}_t \left[\left(\frac{\varepsilon_{jt+1}}{\theta_{jt+1}} \right)^\alpha \right] \right)^{\frac{1}{\alpha}} \quad (1.18)$$

$$Y_{jt+1} = \varepsilon_{jt+1} F^j(K_{jt+1}) \quad (1.19)$$

1.3.3 First-order conditions

The first-order condition for the productivity level ε_{jt+1} in each state of nature is given by (all the algebra is provided in the appendix)

$$\frac{\varepsilon_{jt+1}}{\varepsilon_{jt}} = \eta_{jt}^{\frac{1}{1-\alpha}} \left(\frac{M_{t+1}P_{jt+1}}{P_{jt}} \right)^{\frac{1}{\alpha-1}} \left(\frac{\theta_{jt+1}}{\theta_{jt}} \right)^{\frac{\alpha}{\alpha-1}} (1 - \tau_{jt+1})^{\frac{1}{\alpha-1}}, \quad (1.20)$$

12. Without loss of generality I specify that technology 1 is the numeraire, i.e. $P_{1t} = 1$

13. See Cochrane (2009) chapter 4, Duffie (2010) chapter 1, or Campbell (2017) chapter 4 for further references on the uniqueness of the SDF.

where

$$\eta_{jt} = \frac{\mathbb{E}_t \left[M_{t+1} \left(\frac{P_{jt+1}}{P_{jt}} \right) (1 - \tau_{jt+1}) \right]}{\mathbb{E}_t \left[\left(\frac{\varepsilon_{jt+1}}{\varepsilon_{jt}} \right)^{\alpha-1} \left(\frac{\theta_{jt+1}}{\theta_{jt}} \right)^{-\alpha} \right]}. \quad (1.21)$$

Eq. (1.20) states that the firm's optimal choice of productivity level in each state of nature is determined by prices, technological constraints, and distortions to output. The firm chooses a higher output in states of nature where it is more valuable. Since $\alpha > 1$, such states include high M_{t+1} and P_{jt+1} , low τ_{jt+1} , and high θ_{jt+1} . Since η_{jt} is predetermined at time t , it does not affect excess returns.

Now, solving for the SDF from the producer's maximization choice yields

$$M_{t+1} = \eta_{jt} \left(\frac{P_{jt+1}}{P_{jt}} \right)^{-1} \left(\frac{\varepsilon_{jt+1}}{\varepsilon_{jt}} \right)^{\alpha-1} \left(\frac{\theta_{jt+1}}{\theta_{jt}} \right)^{-\alpha} \frac{1}{(1 - \tau_{jt+1})} \quad (1.22)$$

Note that a constant distortion level would not have a first-order impact on the pricing kernel, although it might have dynamic implications. State-contingent distortions should directly impact the relative prices of contingent claims. Eq. (1.22) states that producer j 's MRT should be equal to the SDF. This is the central equation in Belo (2010), it states that every asset can be valued without considering any information on consumer preferences. Eq. (1.22) can be expressed in terms of output to simplify the empirical implementation

$$M_{t+1} = \bar{\eta}_{jt} \frac{1}{(1 - \tau_{jt+1})} \left(\frac{P_{jt+1}}{P_{jt}} \right)^{-1} \left(\frac{Y_{jt+1}}{Y_{jt}} \right)^{\alpha-1} \theta_{jt+1}^{-\alpha}, \quad (1.23)$$

where

$$\bar{\eta}_{jt} = \frac{\mathbb{E}_t \left[M_{t+1} \left(\frac{P_{jt+1}}{P_{jt}} \right) (1 - \tau_{jt+1}) \right]}{\mathbb{E}_t \left[\left(\frac{Y_{jt+1}}{Y_{jt}} \right)^{\alpha-1} \theta_{jt+1}^{-\alpha} \right]} \quad (1.24)$$

is again a variable pre-determined at time t .

1.3.4 Identification

In order to use the SDF in Eq. (1.23), I need to measure the unobserved underlying productivity level θ_{jt} . Therefore, following Belo (2010), I assume that the underlying productivity level in each technology $j = 1, \dots, N$ has the following factor structure

$$\alpha \log(\theta_{jt}) = \lambda_j \bar{\theta}_t, \quad (1.25)$$

where $\bar{\theta}_t$ is the common (across technologies) productivity factor and λ_j is the loading of the underlying productivity level of technology j on the common productivity factor. Without loss of generality, I normalize the loadings for technology 1 to $\lambda_1 = 1$. The loadings for technology j , λ_j , capture the differences in sensitivity for each technology with respect to the business cycle.

In order to identify the SDF in Eq. (1.23), following the literature in misallocation, I approximate the output distortion in each technology $j = 1, \dots, N$, with the following relationship

$$(1 - \tau_{jt+1}) = \Phi_{jt} \text{TFP}_{jt+1}. \quad (1.26)$$

where TFP_{jt+1} is the total factor productivity (TFP) for technology j at time t . This

assumption is motivated by the documented relationship between productivity and misallocation. As discussed in section 5.2.1 of Bloom et al. (2018), one of the channels of output dispersion is misallocation. In their model, an increase in misallocation acts as a negative first-moment shock to productivity. Misallocation increases in the economy in response to a TFP shock. This is captured by the inverse relationship that I propose between τ_{jt+1} and TFP.

As Restuccia and Rogerson (2008) documented, misallocation can have direct implications on the level of TFP, in the range of 30 to 50 percent. This effect is present in their model when policies create heterogeneity in prices. Finally, Hsieh and Klenow (2009), provide quantitative evidence on the potential impact of resource misallocation on TFP.

Assume Φ_{jt} represents policies made for specific firms that distort next period TFP. The misallocation could load differently through time, as described by Bloom et al. (2018) findings that uncertainty is countercyclical. Restuccia and Rogerson (2008) propose that firms face constant output distortions on the steady state. Here, I generalize this assumption. This represents policies made for specific technologies which smooth or intensify the ripples of the business cycle.

Equations (1.25) and (1.26) impose a restriction between the producers' first-order conditions, which can be used to infer the underlying productivity level factor and the MRT from the observed output, price, and TFP data in the different technologies. This result is stated in Proposition 1.

Proposition 1.

Consider any two technologies, equations (1.25) and (1.26) imply that the common

productivity factor $\bar{\theta}$ and the equilibrium MRT M can be identified from output, price, and TFP data in two technologies. The common productivity factor and the equilibrium MRT are given by

$$\bar{\theta}_{t+1} = \frac{\zeta_{1t} - \zeta_{2t}}{1 - \lambda} - b^p \Delta p_{2t+1} - b^y (\Delta y_{1t+1} - \Delta y_{2t+1}) - b^\tau (\text{tfp}_{1t+1} - \text{tfp}_{2t+1}) \quad (1.27)$$

$$M_{t+1} = \kappa_t \exp \left(b^p \Delta p_{2t+1} + b^y (\lambda \Delta y_{1t+1} - \Delta y_{2t+1}) + b^\tau (\lambda \text{tfp}_{1t+1} - \text{tfp}_{2t+1}) \right) \quad (1.28)$$

where $\zeta_{jt} = \log(\bar{\eta}_{jt}) - \log(\Phi_{jt})$, $\kappa_t = \exp((1 - \lambda)^{-1}(\zeta_{2t} - \lambda \zeta_{1t}))$, $\lambda = \lambda_2$, lower case letters represent logs, and Δ is the difference operator. The factor risk prices b^p , b^y , and b^τ are given by

$$\begin{bmatrix} b^p \\ b^y \\ b^\tau \end{bmatrix} = \frac{1}{1 - \lambda} \begin{bmatrix} -1 \\ -(\alpha - 1) \\ 1 \end{bmatrix} \quad (1.29)$$

Proof.

Taking logs on the SDF in Eq. (1.23) for two arbitrary producers $j = 1, 2$ (in which $j = 1$ is the numeraire good), and evaluating the log MRT for two different technologies gives the result for $\bar{\theta}_{t+1}$ and M_{t+1} .

1.3.5 Asset Pricing Implications

The MRT in Eq. (1.28) is a valid SDF. For a vector of excess returns R_{t+1}^e , a valid SDF satisfies:

$$\mathbb{E}_t \left[M_{t+1} R_{t+1}^e \right] = 0. \quad (1.30)$$

Up to a constant at time t , any valid discount factor satisfies Eq. (1.30). Thus, as a simplification, normalize κ_t to 1, then the SDF becomes,

$$M_{t+1} = \left(\frac{P_{2t+1}}{P_{2t}}\right)^{b^p} \left(\frac{Y_{1t+1}}{Y_{1t}}\right)^{\lambda b^y} \left(\frac{Y_{2t+1}}{Y_{2t}}\right)^{-b^y} (\text{TFP}_{1t+1})^{\lambda b^\tau} (\text{TFP}_{2t+1})^{-b^\tau} \quad (1.31)$$

where P_{2t} are the relative prices of technology 2 with respect to the numeraire. I estimate and test the model using different portfolios of assets. In order to estimate the expected excess return using the SDF from Eq. (1.31), I use the following relationship between excess returns and risk premium

$$\mathbb{E}_t [R_{t+1}^e] = -\frac{\text{Cov}_t (M_{t+1}, R_{t+1}^e)}{\mathbb{E}_t [M_{t+1}]}. \quad (1.32)$$

1.4 Quantitative Analysis

This section presents the quantitative assessment of the model.

1.4.1 Data

I use macroeconomic data for two technologies to numerically estimate the SDF from Eq. (1.31). This specification requires price, output, and total factor productivity data for two production technologies. Following Gomes, Kogan, and Yogo (2009) and Belo (2010), I choose the durables and nondurable goods sectors as technologies 1 and 2, respectively. According to Eq. (1.31), the equilibrium MRT is

$$M_{t+1} = \left(\frac{P_{NDt+1}}{P_{NDt}}\right)^{b^p} \left(\frac{Y_{Dt+1}}{Y_{Dt}}\right)^{\lambda b^y} \left(\frac{Y_{NDt+1}}{Y_{NDt}}\right)^{-b^y} (\text{TFP}_{Dt+1})^{\lambda b^\tau} (\text{TFP}_{NDt+1})^{-b^\tau}, \quad (1.33)$$

where $j = D$ represents durables, $j = ND$ represents nondurables, and the factor risk prices b^p , b^y , and b^r are given in Eq. (1.29).

The selection of durables and nondurables hinges on the time-series differences that we observe in the data, which represents a convenient modeling choice. Table 1.1 shows the loads of durables and nondurables on the underlying production level differ during expansions and recessions.¹⁴ Therefore, durable and nondurable technologies have different business cycle sensitivities. This suggests that durables and nondurables load differently on the common productivity factor (i.e., $\lambda \neq 1$).

Table 1.1
Summary statistics.

	Full sample			NBER expansions		NBER recessions	
	Mean	S.D.	AC(1)	Mean	S.D.	Mean	S.D.
tfp_D	2.75	2.23	0.12	2.79	2.14	2.61	2.66
tfp_{ND}	0.73	1.86	0.10	0.62	1.67	1.23	2.54
Δy_D	4.89	6.35	-0.15	5.70	4.89	1.40	10.08
Δy_{ND}	2.73	2.20	-0.14	2.77	2.15	2.58	2.49
Δp_D	1.45	2.80	0.76	1.24	2.51	2.39	3.78
Δp_{ND}	2.38	2.76	0.48	2.37	2.78	2.44	2.77
m	0.00	1.50	-0.13	-0.30	0.74	1.27	2.83
$\bar{\theta}$	0.00	1.06	-0.11	0.14	0.88	-0.62	1.52

This table shows the mean, standard deviation S.D, and autocorrelation AC(1) for the log total factor productivity tfp_j , output growth Δy_j , and price growth Δp_j , where $j = D$ are durable goods, and $j = ND$ are nondurable goods. It also presents the demeaned log SDF m and the demeaned productivity level $\bar{\theta}$. All values are expressed in percentage points. The data are annual and the sample covers 1948-2021.

The macro-data for output and prices comes from the National Income Product Accounts (NIPA) available through the Bureau of Economic Analysis (BEA) website.

14. A year is defined as a recession year if at least five months in that year are defined as being in a recession by the NBER.

Output in each industry is measured by the real gross domestic product (Table 1.2.3, lines 7 and 10). Price data in each industry is measured by the GDP price deflator (Table 1.2.4, lines 7, 10). TFP is the annual U.S. Solow residual for durables (TFP in equipment and consumer durables) and nondurables (TFP in non-equipment business output ("consumption")), downloaded from John Fernald's website (Fernald 2014). The data is annual and the sample covers 1948-2021.

As we can observe from Table 1.1, output growth is cyclical regardless of the technology, but TFP and price growth display different behavior for durables and nondurables goods. Even though output growth is cyclical for the two technologies, the sensitivity of each industry to the business cycle differs. The difference in output growth for durables between expansions and recessions is 4.30% whereas for non-durables is only 0.19%. Additionally, price growth for nondurables is acyclical with a difference of only 0.07%, and durables are countercyclical with 1.15% lower during expansions. Contrary to output growth and price growth, the TFPs display larger differences for nondurables across expansions and recessions. Finally, as expected, the MRT is countercyclical, and the technology level is cyclical.

1.4.2 Estimation

I use the generalized method of moments (GMM), to estimate the parameters α and λ . I follow the methodology developed by Hansen and Singleton (1982). The following moment restriction is used for estimation and testing:

$$\mathbb{E} \left[M_{t+1} R_{it+1}^e z_t \right] = 0 \tag{1.34}$$

where R_{it+1}^e is the excess return for test portfolio $i = 1, \dots, N_{tp}$ and z_t corresponds with the instruments used for the GMM estimation.¹⁵ The instruments condition the information until time t . First, in Section 1.4.3, I use the 6 portfolios sorted on size and book-to-market as main test assets. In Section 1.4.4, I present additional test portfolios and obtain similar results. The value-weighted excess return for the test asset comes from Kenneth French’s webpage. The instruments are a constant and the dividend-price ratio.¹⁶ The data for the dividend-price ratio come from Robert Shiller’s webpage. Additionally, the asset return at time t is matched with the macro-variable at time $t + 1$ (Campbell 2003).

The number of moment conditions is $N_{tp} \times N_i$ where $N_i = 2$ is the number of instruments. The number of parameters to estimate is $N_p = 2$ (α and λ). Therefore, this setup is over-identified with $N_{tp} \times N_i - N_p$ overidentifying restrictions. I use the J -test to see if the overidentified restrictions are statistically different from zero. I report the two GMM stages, in the first stage, the weighting matrix is the identity matrix whereas, for the second stage, I use the efficient weighting matrix corresponding to the inverse of the Newey-West estimate of the sample pricing errors’ covariance matrix. Following Belo (2010), I use two-period lags for the Newey-West estimate of the sample pricing errors’ covariance matrix to account for the possibility of time aggregation in output and price data.¹⁷ In addition to testing if the sample pricing errors’ are identically zero using the J -test, I measure the cross-sectional R-squared (R^2) and the mean absolute pricing error (MAPE). The R^2 is obtained from an OLS

15. See Cochrane (2009) Chapters 10 and 11 for further details of the GMM methodology.

16. I normalize to 1 the dividend-price ratio mean, as suggested in Cochrane (2009).

17. Hall (1988) discusses issues with the time aggregation of consumption data. In the context of this study, similar issues may arise with the output, price, and TFP data.

regression of the realized average excess return of each portfolio on the average expected excess return of each portfolio. I estimate the average excess return of each portfolio using Eq. (1.32). To estimate the MAPE, first I compute the absolute pricing error of each portfolio $\text{Pricing Error}_i = |\mathbb{E}[R_i^e]^{\text{observed}} - \mathbb{E}[R_i^e]^{\text{predicted}}|$, and then averaging them $\text{MAPE} = (\sum_{i=1}^{N_{tp}} \text{Pricing Error}_i) / N_{tp}$.

1.4.3 Results

Table 1.2 summarizes the main findings of this study. The first and second columns show the results for the GMM's first and second stages, respectively. For comparisons, I also perform the same analysis without the inclusion of misallocation (standard Belo model), these results are in columns 3 and 4 for the GMM first and second stages respectively. When considering misallocation, the firm's flexibility to adapt across states is severely reduced, measured by the curvature parameter α . The point estimate of the curvature parameter α of 3.80 is more than twice compared to 1.66 without misallocation. The two point estimates differ considerably with Belo (2010) findings. He finds a curvature parameter α close to 1, which means that firms are adaptable to every state. We can attribute an increase from Belo's α of 1.02, to the α of 1.66 I estimate, to the differences in the sample periods.¹⁸ Therefore, if I consider a baseline α estimate of 1.66 with no misallocation, the increase of this value to 3.80 is caused by firm's distortions. Misallocation not only reduces the firm's output as observed by the firm's maximization problem, but also affects the ability the firm has to adapt across states of nature. A larger α implies that firm's

18. Belo (2010) uses annual data and his sample period covers 1930-2007. I also estimate Belo's model using his time period, and find almost identical results to the parameters α and λ found in Section 3.5 of Belo (2010).

chosen productivity level ε has to closely follow θ to satisfy the productivity-choice constraint in Eq. (1.12). This suggests the difference between the two α provides an indirect measure of the extent to which firms are effectively transforming output across states.

Table 1.2
GMM estimation of the production-based model.

	With Misallocation		No Misallocation	
	1st	2nd	1st	2nd
α	4.79	3.80	2.27	1.66
s.e.	16.57	5.96	3.85	0.31
λ	0.82	0.85	0.93	0.96
s.e.	0.67	0.27	0.20	0.02
R^2		81.06		85.50
MAPE		1.08		1.33
J -test	15.30	15.81	14.54	15.08
p -Value (J)	12.17	10.51	14.99	12.93
M.C.	12.00	12.00	12.00	12.00

This table shows the results for the GMM first-stage and second-stage. The moment conditions are $\mathbb{E}[M_{t+1}R_{it+1}^e z_t] = 0$ where M_{t+1} is the SDF, z_t are the instrumental variables, and R_{it+1}^e are the excess returns of the six size and book-to-market portfolios. Columns 1 and 2 present the results for the SDF in Eq. (1.31), and columns 3 and 4 use the same specification but with no misallocation. The estimation for the two stages is presented for α , λ , and the J -test for the overidentifying restrictions. R^2 corresponds to the regression of the mean observed excess return on the predicted expected excess return. MAPE is the mean absolute value of the difference between the mean observed excess return and the mean expected excess return. M.C. corresponds to the number of moment conditions. R^2 , MAPE, and p -Value (J) are in percentage points. The data is annual and the sample covers 1948-2021.

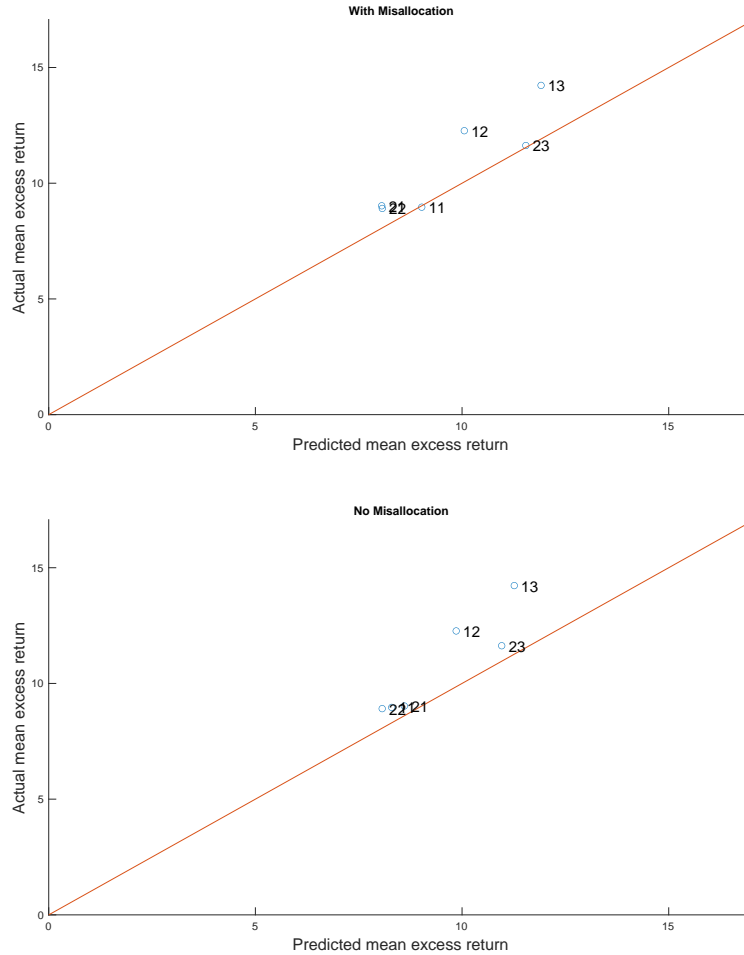
The parameter λ measures the sensitivity of the underlying productivity level in the nondurable goods sector to the common productivity factor. I find a λ of 0.96 when no distortions are present. This result is almost identical to Belo's finding of λ of 0.97. This indicates the nondurables sector is slightly less sensitive to variations

in the common productivity factor $\bar{\theta}$ than durables. When distortions are taken into account, I find a larger difference, with a point estimate for λ of 0.85. The presence of misallocation reduces the sensitivity of the underlying productivity level in the nondurables to the common factor. A smaller λ can also help to explain why the nondurable goods sector is less cyclical than the output growth in the durable goods sector: nondurables are less sensitive to the common productivity factor $\bar{\theta}$ which is closely related to the business cycle. The parameter λ is also present in the factor risk prices in Eq. (1.29). A smaller λ effectively reduces the magnitude of the prices of risk for durables associated with output growth λb^y and misallocation λb^r , which are cyclical. This cause a tapered effect on the SDF for durables in the business cycle with respect to the model without misallocation.

The J -test of overidentifying restrictions fails to reject the model in both stages, with a p -value of 12.2% and 10.5% for the first and second stages respectively. The model captures well the cross-sectional variation with R^2 of 81.1%, but this value is lower when compared to the original model without misallocation that obtains an R^2 of 85.5%.

The original Belo model also measures the pricing errors in the cross-section. He finds a mean absolute pricing error MAPE of 1.15, which indicates his model does a good job predicting the cross-section of the 6 size and book-to-market portfolios. Table 1.2 shows a MAPE of 1.33 for the model with no misallocation, showing the cross-sectional prediction is reduced in this new sample. In the model with misallocation, MAPE is reduced by 25 basis points (bps) when compared to the model without misallocation. This suggests for the 6 size and book-to-market portfolios misallocation is slightly priced in the cross-section.

Figure 1.2
Pricing Errors



This figure shows the plot of the predicted versus realized excess returns per annum implied by the estimation of the production-based model on the 6 Portfolios Formed on Size and Book-to-Market (2x3). The first digit refers to the size sort (1 and 2 for small and big, respectively), and the second digit refers to book to market sort (1, 2, and 3 indicating the growth, neutral, and value portfolios, respectively). (i) top: considering misallocation; and (ii) bottom: without misallocation (Belo's model). The data is annual and the sample covers 1948-2021.

Figure 1.2 presents a visual explanation of the fit of the production-based model on the 6 size and book-to-market portfolios. This figure plots the predicted versus

realized excess returns implied by the second-stage GMM estimates of the model. The top panel presents the model I study, which explains 81.1% of the variation in average returns in the cross-section, with a mean absolute pricing error of 1.08% per annum. The straight line is the 45-degree line, and it is where the portfolios should lie. In the figure, the first digit represents the size (1 small and 2 big) and the second is the book-to-market (1 growth, 2 neutral, and 3 value). All the portfolios seem to lie on the 45-degree line, with some small pricing errors, especially for small-size portfolios. The bottom panel presents the analysis without misallocation. Similarly, the model without misallocation seems to capture well big-size portfolios but presents small pricing errors for small-size portfolios.

As pointed out by Campbell (2017), the initial time period in Belo’s model has a one-factor structure, where CAPM works well. However, the sample I use in this study from 1948 to 2021 doesn’t hinge on that period, and both models seem to predict observed mean excess return.

1.4.4 Results: Additional Test Assets

In this section, I test additional portfolios of assets. The main goal is to provide a more general view of the results from the previous section. The test portfolios are sorted on: size (ME), book-to-market (BM), operating profitability (OP), investment (IN), momentum (MO), short-term reversal (SR), and long-term reversal (LR). I also include industry portfolios. I use for univariate sort 10 portfolios; for bivariate sort 6 (2x3) and 25 (5x5) portfolios; and for three-way sort 32 (2x4x4) portfolios. For the industry portfolios, I use 5, 10, 12, 17, 30, 38, 48, and 49 portfolios. Finally, I create a 50-asset portfolio formed on the univariate sorting of 10 portfolios for size,

book-to-market, operating profitability, investment, and momentum.

Table 1.3 presents the GMM estimation results for α and λ , R^2 , and MAPE for 33 additional test portfolios. I use the same instruments as in the previous section, a constant and dividend-price ratio. I perform the analysis using the SDF from Eq. (1.33) (columns 2-5), and as a reference, the estimates using the standard Belo (2010) model (columns 6-9). In column 1, we can observe that α is consistent across the assets with a cross-sectional mean of 3.79 and a standard deviation of 0.57. These values contrast with the lower estimates when misallocation is not considered, with a cross-sectional mean of 1.99 with a standard deviation of 0.52. For all the assets tested, the addition of misallocation increases the parameter α . This suggests that misallocation effectively reduces the firm's ability to choose technologies to adapt to exogenous shocks. The loading on nondurables measured by λ is also reduced when misallocation is considered in the model. The cross-sectional average loading λ is 0.84 (s.d. of 0.02) with misallocation, and 0.94 (s.d. of 0.03) for the standard Belo model. This indicates that misallocation reduces the nondurable sensitivity to the underlying technology factor. R^2 and MAPE present a larger dispersion with values that range from 2.8% to 81.2% for R^2 and 1.0% to 13.9% for MAPE for the model with misallocation, and 0.1% to 85.5% for R^2 and 1.0% to 13.5% for MAPE for the model without misallocation. The larger pricing errors are for the portfolios formed on momentum and for the industry portfolios, especially for more than 30 industries. However, the model with misallocation consistently reduces the pricing errors, with a cross-section average reduction of 7 bps.

The main takeaway of table 1.3 is the consistency of the curvature parameter α and the loading on nondurables λ in the cross-section for the two models. This suggests that misallocation has a detrimental effect on the firm's ability to switch

Table 1.3
Additional Test Portfolio Estimation

Portfolio	M.C.	With Misallocation				No Misallocation			
		α	λ	R^2	MAPE	α	λ	R^2	MAPE
6 Portfolios Formed on ME and BM	12	3.80	0.85***	81.06	1.08	1.66***	0.96***	85.50	1.33
6 Portfolios Formed on ME and OP	12	3.71	0.83**	12.75	3.24	1.37***	0.97***	0.43	3.34
6 Portfolios Formed on ME and IN	12	4.29	0.86***	35.85	1.87	2.20**	0.94***	0.06	2.23
6 Portfolios Formed on ME and MO	12	3.53	0.83**	9.86	4.56	1.49***	0.96***	19.45	4.55
6 Portfolios Formed on ME and SR	12	3.34	0.84**	58.53	3.56	2.18*	0.91***	35.54	3.48
6 Portfolios Formed on ME and LR	12	3.81	0.85***	78.16	1.78	1.90**	0.94***	72.05	1.91
10 Portfolios Formed on ME	20	4.26	0.80***	81.19	2.72	2.63*	0.89***	80.28	2.32
10 Portfolios Formed on BM	20	3.54	0.85***	55.84	2.38	1.67***	0.95***	55.84	2.41
10 Portfolios Formed on OP	20	3.73	0.85***	16.14	1.63	1.46***	0.97***	11.36	1.37
10 Portfolios Formed on IN	20	3.74	0.86***	10.37	1.46	1.90***	0.95***	29.67	1.71
10 Portfolios Formed on MO	20	3.31	0.85***	50.02	4.13	1.64***	0.95***	57.32	4.00
10 Portfolios Formed on SR	20	3.73	0.86***	50.05	1.19	2.52**	0.91***	44.02	1.27
10 Portfolios Formed on LR	20	3.44	0.86***	75.72	1.70	2.02***	0.94***	75.70	1.69
25 Portfolios Formed on ME and BM	50	4.19*	0.85***	47.06	1.45	1.97***	0.95***	50.61	1.45
25 Portfolios Formed on ME and OP	50	4.21***	0.86***	2.81	1.95	2.07***	0.95***	5.39	1.90
25 Portfolios Formed on ME and IN	50	4.64***	0.85***	16.52	1.63	2.66***	0.92***	4.30	1.94
25 Portfolios Formed on ME and MO	50	3.53*	0.85***	7.96	4.37	1.91***	0.94***	9.82	4.24
25 Portfolios Formed on ME and SR	50	4.08*	0.85***	36.04	2.29	2.55***	0.91***	23.46	2.53
25 Portfolios Formed on ME and LR	50	4.09**	0.86***	57.19	0.97	2.18***	0.94***	54.37	0.97
25 Portfolios Formed on BM and OP	50	4.25**	0.85***	30.36	1.82	2.08***	0.94***	28.63	1.85
25 Portfolios Formed on BM and IN	50	4.17**	0.86***	24.82	1.39	2.23***	0.94***	17.46	1.63
25 Portfolios Formed on OP and IN	50	4.02**	0.86***	5.48	2.00	2.13***	0.94***	8.97	2.10
32 Portfolios Formed on ME, BM, and OP	64	4.87***	0.83***	18.44	2.86	3.07***	0.90***	14.21	2.75
32 Portfolios Formed on ME, BM, and IN	64	4.81***	0.84***	21.99	2.11	3.36***	0.89***	17.15	2.21
32 Portfolios Formed on ME, OP, and IN	64	4.55***	0.84***	5.01	2.53	2.59***	0.92***	2.15	2.63
5 Industry Portfolios	10	2.90	0.87***	47.08	3.48	1.18***	0.97***	39.59	5.55
10 Industry Portfolios	20	3.70	0.83***	21.40	3.25	1.23***	0.98***	65.55	3.16
12 Industry Portfolios	24	3.71	0.84***	14.67	3.00	1.23***	0.98***	64.43	1.58
17 Industry Portfolios	34	3.38*	0.86***	13.97	2.37	1.50***	0.97***	33.99	1.82
30 Industry Portfolios	60	3.58***	0.86***	4.50	2.48	1.58***	0.96***	14.03	2.13
38 Industry Portfolios	76	2.41***	0.74***	27.35	13.89	1.38***	0.94***	2.29	13.53
48 Industry Portfolios	96	2.82	0.81	13.81	7.64	1.92	0.85	7.54	8.89
49 Industry Portfolios	98	2.69***	0.83***	15.47	7.73	2.04***	0.87***	6.92	8.30
50 Portfolios Formed on ME + BM + OP + IN + MO	100	4.08***	0.85***	3.44	1.91	2.03***	0.94***	4.51	2.01

The portfolios are as follows: ME represents size, BM is book-to-market, OP is operating profitability, IN is investment, MO is momentum, SR is short-term reversal, and LR is long-term reversal. This table shows the results for the GMM estimation of α and λ . R^2 corresponds to the regression of the mean observed excess return on the predicted expected excess return. MAPE is the mean absolute value of the difference between the mean observed excess return and the mean expected excess return. M.C. corresponds to the number of moment conditions. Significance is represented by * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. R^2 and MAPE values are in percentage. The data is annual and the sample covers 1948-2021, with the exception of the portfolios formed on OP or IN where the sample covers 1964-2021.

output across states.

1.5 Simulation Analysis

1.5.1 Explicit Representation of a Single Firm

I solve the model for a single representative firm with Cobb-Douglas production in a competitive economy. The firm faces two states of nature, a high state $s = h$ and a low state $s = l$. Past states are correlated with future states. This is relevant for the underlying productivity distribution from where the firm draws ε which depends on the previous and current period. This can be represented as follows

$$\begin{aligned} 1 &= \left(\mathbb{E}_t \left[\left(\frac{\varepsilon(s_{t+1})}{\theta(s_{t+1})} \right)^\alpha \right] \right)^{\frac{1}{\alpha}} \\ &= \left(p_{s_{t+1}=h|s_t} \left(\frac{\varepsilon(s_{t+1}=h|s_t)}{\theta(s_{t+1}=h)} \right)^\alpha + p_{s_{t+1}=l|s_t} \left(\frac{\varepsilon(s_{t+1}=l|s_t)}{\theta(s_{t+1}=l)} \right)^\alpha \right)^{\frac{1}{\alpha}}, \end{aligned} \quad (1.35)$$

where $s_{t+1} = h|s_t$ ($s_{t+1} = l|s_t$) is the transition from the current state s_t to the next state $s_{t+1} = h$ ($s_{t+1} = l$). Note that given the feasibility constraint in Eq. (1.35), specifying the high state and the transition probabilities, fully specifies the low state

$$\varepsilon(s_{t+1} = l|s_t) = \left[\frac{1}{p_{s_{t+1}=l|s_t}} - \frac{p_{s_{t+1}=h|s_t}}{p_{s_{t+1}=l|s_t}} \left(\frac{\varepsilon(s_{t+1}=h|s_t)}{\theta(s_{t+1}=h)} \right)^\alpha \right]^{\frac{1}{\alpha}} \theta(s_{t+1}=l) \quad (1.36)$$

The optimal productivity level ε in the low state follows directly from other known or chosen quantities.

For further tractability, the firm produces a numeraire good with a price normalized to 1. Writing the dynamic problem in Bellman form allows for the application of a traditional discrete state-space solution using the techniques described in Ljungqvist and Sargent (2012):

$$V(K, \varepsilon(\bar{s}), \bar{s}) = \max_{\varepsilon(\bar{s}), K_{t+1}} \{D(\bar{s}) + \mathbb{E}[M(\bar{s}') V(K', \varepsilon(\bar{s}'), \bar{s}')]\} \quad (1.37)$$

where

$$D(\bar{s}) = Y(\bar{s})(1 - \tau(s)) - (K' - (1 - \delta)K) \quad (1.38)$$

$$Y(\bar{s}) = \varepsilon(\bar{s}) K^\gamma \quad (1.39)$$

$$\varepsilon(\bar{s}) \equiv \begin{cases} \varepsilon(h|s_{-1}) & s = h \\ \left[\frac{1}{p_{l|s_{-1}}} - \frac{p_{h|s_{-1}}}{p_{l|s_{-1}}} \left(\frac{\varepsilon(h|s_{-1})}{\theta(h)} \right)^\alpha \right]^{\frac{1}{\alpha}} \theta(l) & s = l \end{cases} \quad (1.40)$$

where to simplify the notation, I define the state tuple $\bar{s} = (s_{-1}, s)$, where s_{-1} represents the previous state, and s the current state. For example, if the previous state was $s_{-1} = l$ and the current state is $s = h$, the state tuple $\bar{s} = (l, h)$. As standard in the macroeconomic literature, prime variables denote the next period value.

Solving the model requires an explicit representation of the pricing kernel. The previous definition for a single firm and numeraire prices is:

$$M(\bar{s}') = \bar{\eta} \left(\frac{1}{1 - \tau(s')} \right) \left(\frac{Y(\bar{s}')}{Y(\bar{s})} \right)^{\alpha-1} \left(\frac{\theta(s')}{\theta(s)} \right)^{-\alpha} \quad (1.41)$$

where

$$\bar{\eta} = \left(\mathbb{E} \left[\frac{1}{(1 - \tau(s'))} \left(\frac{Y(\bar{s}')}{Y(\bar{s})} \right)^{\alpha-1} \left(\frac{\theta(s')}{\theta(s)} \right)^{-\alpha} \left(\gamma \frac{Y(\bar{s}')}{K'} (1 - \tau(s')) + (1 - \delta) \right) \right] \right)^{-1} \quad (1.42)$$

Finally, this result must be de-trended in order to generate a finite and tractable solution:

$$\tilde{M}(\bar{s}') = \beta \times M(\bar{s}') \times \mathbb{E}[M]^{-1} \quad (1.43)$$

$$= \beta \times M(\bar{s}') \quad (1.44)$$

where I normalized the risk-free rate to 1.¹⁹ Note that this sacrifices the model's absolute interpretation. However, the purpose of the derivation is to determine the relative impact of state-contingent prices on relative output choices. These interpretations remain valid.

The transition matrix is calculated given NBER data for the average duration of recessions (10.3 months) and booms (68.8 months). Inverting this gives the following approximate transition matrix, where the $p_{i,j}$ element corresponds to the probability of transitioning from state i to state j

$$P = \begin{bmatrix} 0.84 & 0.16 \\ 0.71 & 0.29 \end{bmatrix}. \quad (1.45)$$

Hence the dynamic programming problem is fully specified.

19. Cochrane (2021) recommends such assumptions for numerical tractability.

1.5.2 Numerical Approach

I calibrate the model at an annual frequency and report calibrated parameters in Table 1.4. The subjective discount factor β and the rate of capital depreciation are obtained from Bloom et al. (2018). I calibrate the productivity distribution ε in high states of nature following Bloom et al. (2018).²⁰ The productivity distribution depends on the previous state, thus there are two possible values for ε in the high state. I normalize ε to 1 for the case of $\bar{s} = (h, h)$. I calibrate the productivity distribution ε for $\bar{s} = (l, h)$ to match their second-moment macroproductivity shocks. Their specification fits in this model given the similarities in the dynamics of uncertainty, which increases during recessions. The α parameter and productivity factor $\bar{\theta}$ were estimated in the previous section for the 6 portfolios formed on size and book-to-market.

My main results use 1000x1000 levels for the continuous variable and 4 levels for the discrete variable. The code used is available upon request.

1.5.3 Scenario Analysis

Consider four different misallocation scenarios. To avoid ambiguities, I label distortions as tax when τ is positive and call them subsidy when τ is negative.²¹ The baseline case considers no distortions. In the second scenario, firms face a 40% tax, labeled as “tax40”, in the high state and no distortion in the low state. In the third

20. The productivity distribution ε in low states is determined by the high state distribution as shown in Eq. (1.40)

21. The confusion may arise given that when the distortion parameter τ is positive (negative) it has a negative (positive) impact on the firm.

Table 1.4
Calibrated and Estimated Parameter Values

Parameter	Value	Description
α	3.80	Curvature, transformation ability across states (estimated)
β	0.95	Time preference parameter (Bloom et al. 2018)
γ	0.30	Returns to scale to capital (following the literature)
δ	0.10	Depreciation (Bloom et al. 2018)
$\theta(h)$	1.15	Productivity, high state (estimated)
$\theta(l)$	0.54	Productivity, low state (estimated)
$\varepsilon(h, h)$	1.00	Productivity, from high to high state (reference)
$\varepsilon(l, h)$	0.98	Productivity, from low to high state (relative) (Bloom et al. 2018)
τ	-	Output distortion, scenario dependent

This table reports the parameter values of the model which are calibrated at an annual frequency.

scenario, the firm receives a subsidy of 40% of production when operating in the low state. Call this specification the “subsidy40” scenario. In the final scenario, firms receive a 20% tax in the high state and a subsidy of 20% in the low state. While not revenue neutral, this scenario is at least revenue hedged as compared with the subsidy40 or tax40 scenarios. Label this the “tax20subsidy20” scenario. The notation and scenarios are summarized in table 1.5.

Table 1.5
Misallocation Scenario Values

Scenario Name	$\tau(h)$	$\tau(l)$
baseline	0%	0%
tax40	40%	0%
subsidy40	0%	−40%
tax20subsidy20	20%	−20%

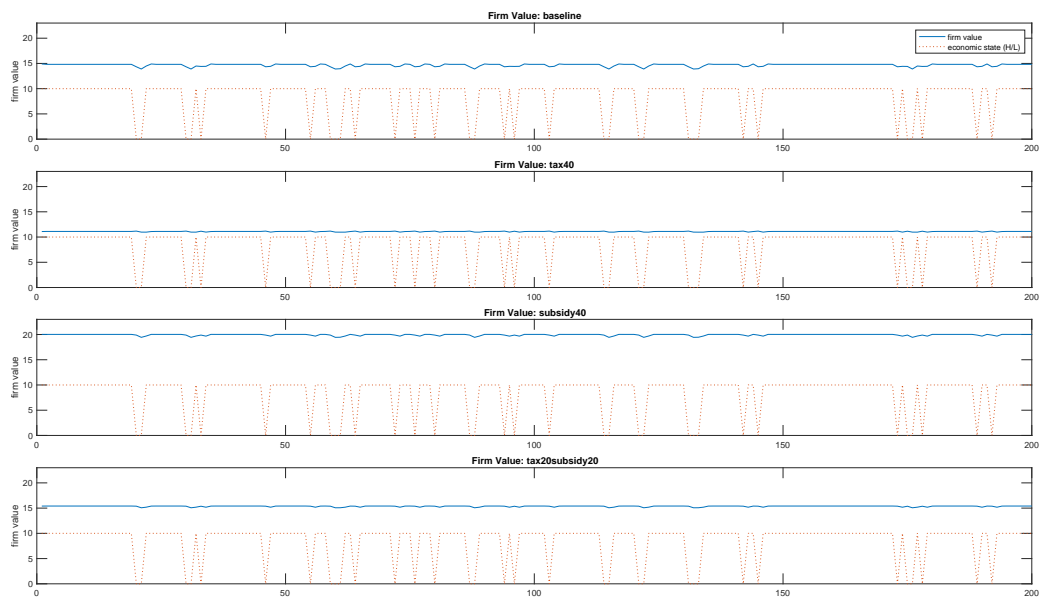
Clearly, the subsidized firm will have a valuation higher than the negatively

distorted firm, yet the overall relative production allocations are not obvious ex-ante. Figure 1.3 shows the predicted effect on value. Such a result is obvious ex-ante but serves to validate the model ex-post. The figure also shows that the value of the firm remains relatively stable over time and is not overly sensitive to the state of the world.

Figure 1.4 displays the effect of the misallocation scenarios on the production allocation. The representative firm keeps a constant level of production during high states and immediately switches to a lower level when the state is low, presenting a cyclical behavior. The model predictions are in line with the simulated results: output is cyclical and is larger when distortions are lower. The average level of output during high states is higher (lower) in the baseline scenario compared to the average level of output during high states when taxes (subsidies) are present. Therefore, misallocations have a direct impact on output, increasing or reducing the production level when firms face subsidies or taxes respectively. Figure 1.5 shows capital that follows a similar pattern to output (cyclical) across all the scenarios.

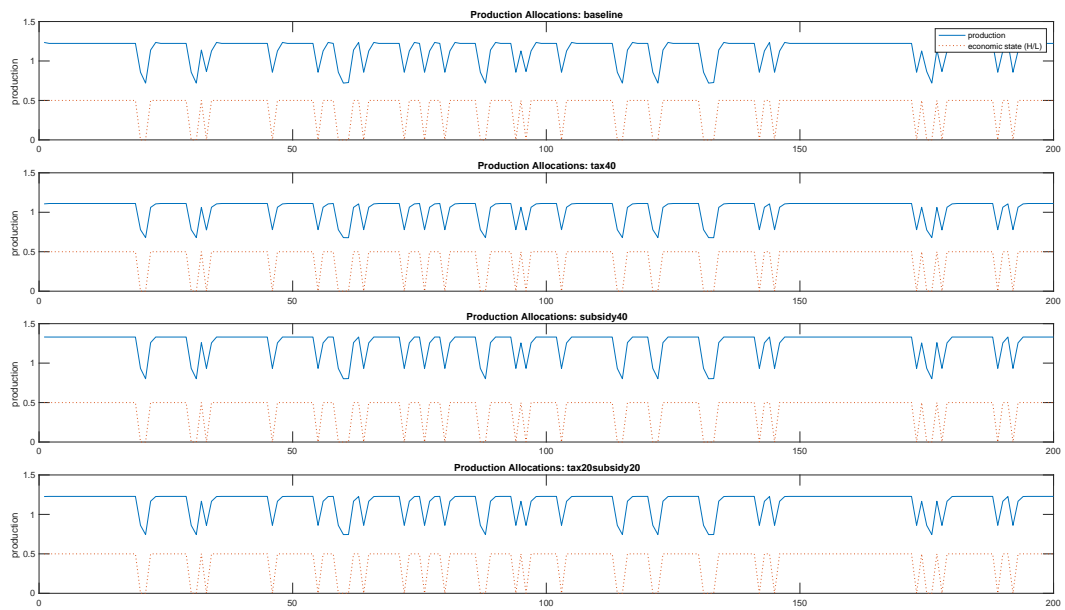
Figure 1.6 and figure 1.7 display the effect of the misallocation scenarios on the firm's investment and dividend. As expected, in the baseline scenario, cyclical behavior is present for investment and dividends. At the through of the business cycle, the firm increases its investment to the point at which capital reaches the steady high state. During the increase in investment, the expected investment return peaks, thus the firm reduces its dividends. This can be observed in Figure 1.8, which displays the expected investment return. When the firm reaches the steady state, the expected investment return reaches its lowest point and the firm in response increases its dividend. An interesting behavior occurs in the "tax40" scenario. Distortions lead to a more tapered reduction in investment and production in low states compared to

Figure 1.3
Firm's Value



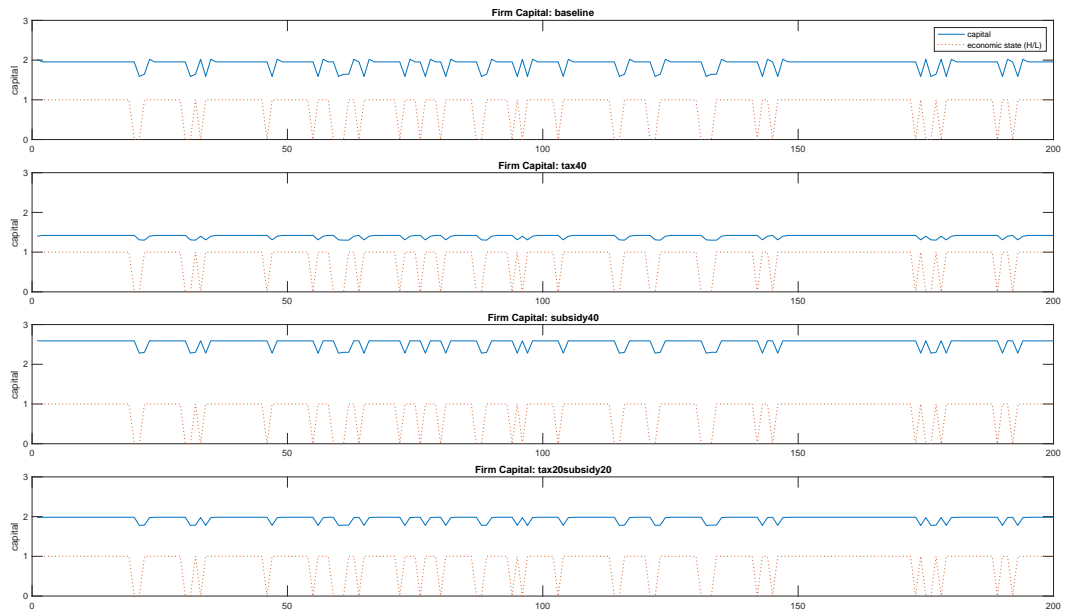
This graph shows the value of the firm across four different misallocation scenarios. In this analysis, all negative distortions are contingent on the realization of a high state, while all positive distortions rely on a low state realization. The actual state realization is represented by the dotted line.

Figure 1.4
Output



This graph shows the firm's output across four different misallocation scenarios. In this analysis, all negative distortions are contingent on the realization of a high state, while all positive distortions rely on a low state realization. The actual state realization is represented by the dotted line.

Figure 1.5
Capital



This graph shows the firm's capital levels across four different misallocation scenarios. In this analysis, all negative distortions are contingent on the realization of a high state, while all positive distortions rely on a low state realization. The actual state realization is represented by the dotted line.

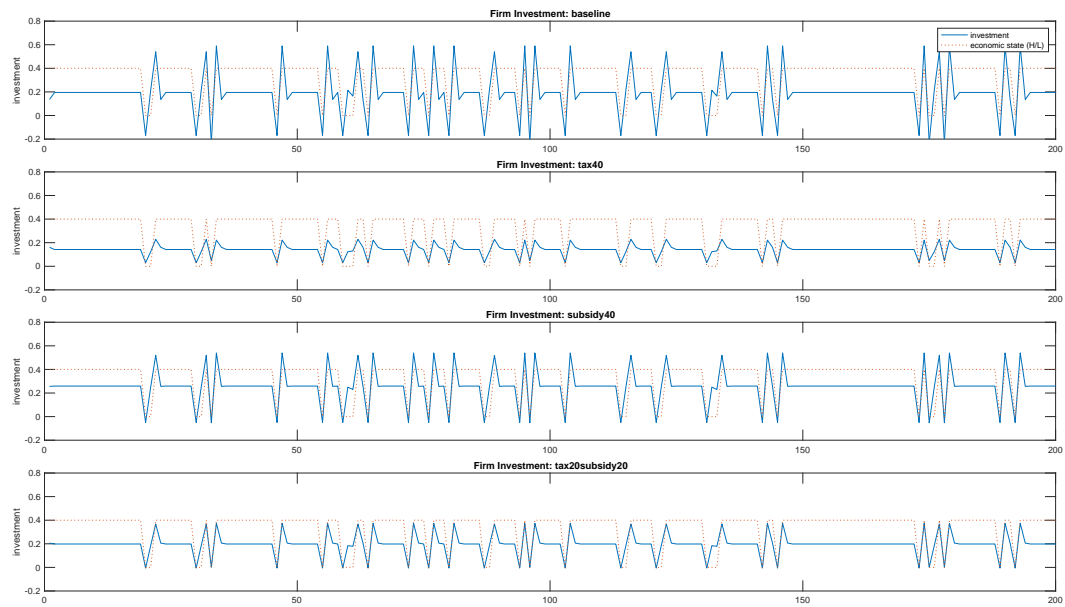
the baseline scenario. When the state transition into a recession, the firm lowers its capital stock reducing its investment. Similarly to the baseline scenario, in the through of the business cycle the investment return peaks, but is much moderated compared to the scenario without distortions. When the investment return peaks, also the return on equity does. The absence of distortions during low-states results in a lower capital level, making it more attractive to increase dividends instead of investment. Misallocations lower the investment return making it more attractive to increase dividends than investing in capital stock. The “subsidy40” scenario shows that the firm accumulates more capital stock than in the baseline scenario during expansions and recessions. The expected investment return is lower compared to the baseline scenario. During the business cycle through, the firm has to reduce its investment to the steady state during recessions. This results in selling its capital, and increasing dividends. Similar to the two previous misallocation scenarios, in the “tax20subsidy20” scenario, misallocations generate a more tapered cyclical investment, but countercyclical dividends, with a lower level of investment return compared to the baseline scenario.

The results indicate that misallocation has a detrimental impact on firms, especially on investment. The expected investment return is much lower when firms face misallocation, making them prioritize dividends instead of investment.

1.5.4 Comparative Statics of Contingent Distortions

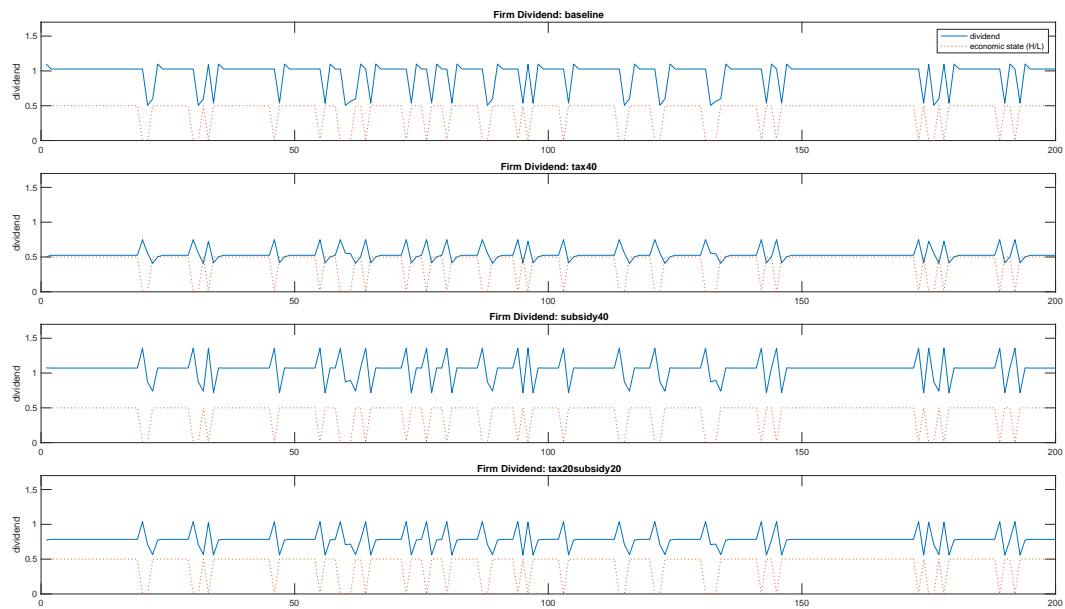
This section serves to reinforce the previous scenario-based approach. The conclusions largely follow from the previous section but cover a wider range of potential state-contingent misallocations. Specifically, I graph the mean realizations or choices

Figure 1.6
Investment



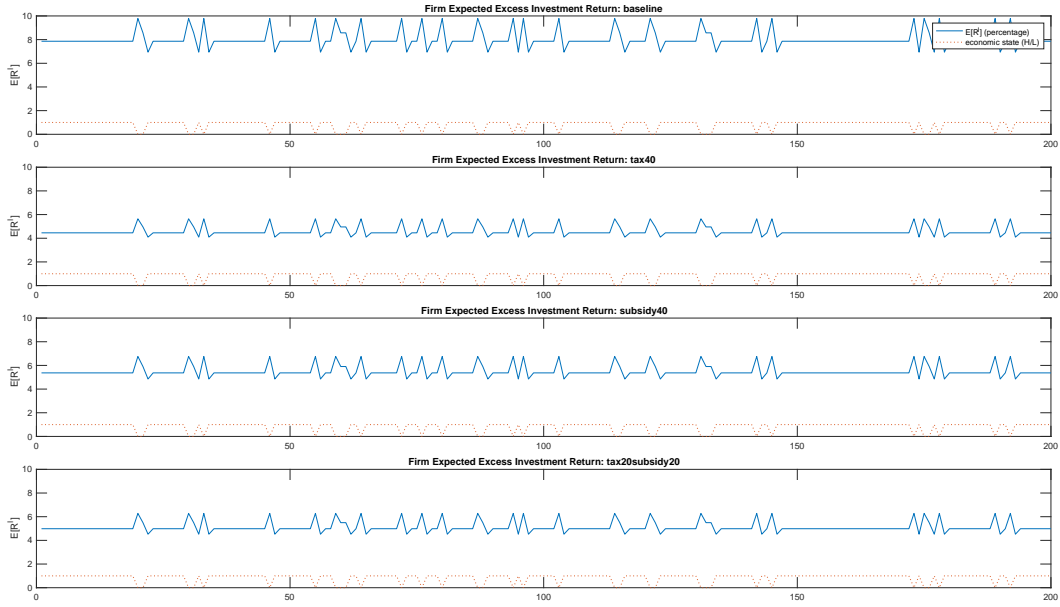
This graph shows the firm's investment levels across four different misallocation scenarios. In this analysis, all negative distortions are contingent on the realization of a high state, while all positive distortions rely on a low state realization. The actual state realization is represented by the dotted line.

Figure 1.7
Dividend



This graph shows the firm's dividend levels across four different misallocation scenarios. In this analysis, all negative distortions are contingent on the realization of a high state, while all positive distortions rely on a low state realization. The actual state realization is represented by the dotted line.

Figure 1.8
Expected Excess Investment Return



This graph shows the firm's expected excess investment return in percentage points across four different misallocation scenarios. In this analysis, all negative distortions are contingent on the realization of a high state, while all positive distortions rely on a low state realization. The actual state realization is represented by the dotted line.

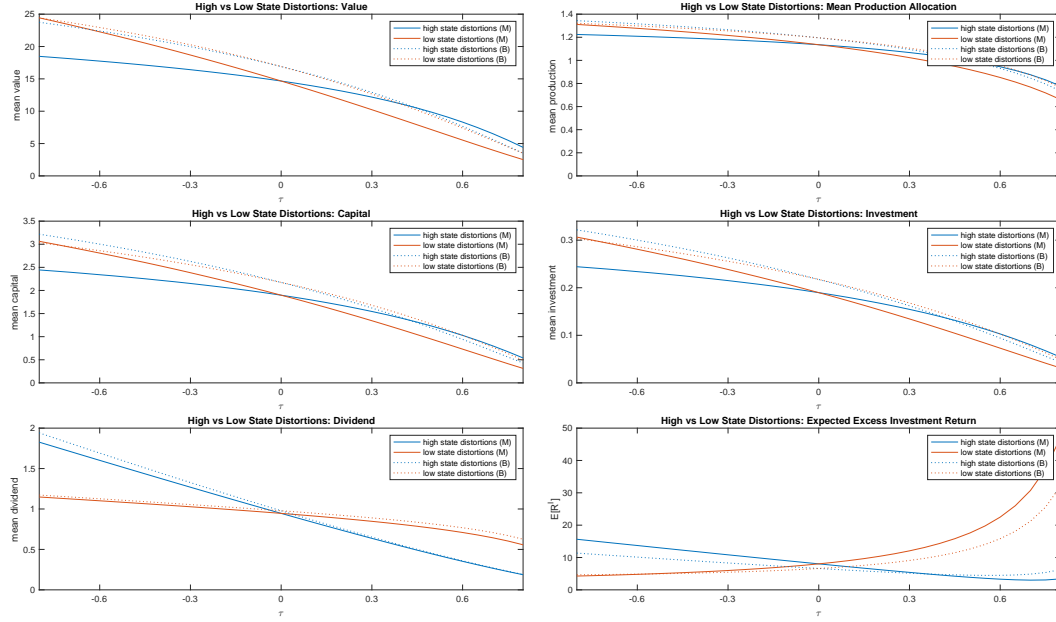
of firm value, production, capital, investment, dividends, and the expected investment return, versus the distortion level of a state of nature. Misallocations in the other state are held at 0. Misallocations range from -0.8 (subsidies) to 0.8 (taxes).

Figure 1.9 shows the effect of varying misallocation in high-states (solid-blue) and low-states (solid-red). Firm value, production, capital, investment, and dividend monotonically decrease with the increase of τ (from subsidies to taxes). With subsidies (negative τ region), firm's steady state is higher for value, production, capital, and investment in the low state (recessions) compared to the scenario without distortions ($\tau = 0$). Since the firm is above its optimal level of investment and

capital stock when there is no misallocation, the investment return is lower. The firm, therefore, increases its value but at the cost of the subsidies instead of a more efficient capital allocation. During high states (expansions), subsidies have a more tapered effect on all the variables. There is a small increase in expected investment return, but since the expected investment return during low states is comparatively much lower (recall that the subsidies are held at zero in low states), the firm does not substantially increase investment but instead increases dividends. The overall effect of subsidies is an inefficient allocation of resources in both states, incentivizing dividends during booms, and higher but less efficient production levels during recessions. With taxes (positive τ region), firm's steady state is lower for almost all variables regardless of the state compared to the scenario without distortions ($\tau = 0$). The only exception is the investment return during recessions, which is almost five times the value when $\tau = 0$, as a result of the extremely low levels of production during recessions. Figure 1.9 shows that distortions have a detrimental overall effect, with larger inefficiencies during low-states (recessions).

Figure 1.9 also shows the effect of estimating the model using the curvature parameter α of 3.80 estimated considering misallocation (solid lines), and α of 1.66 without misallocation (dotted lines). The previous analysis holds for $\alpha = 1.66$, firm value, production, capital, investment, and dividend monotonically decrease with the increase of misallocation τ . However, firm value is less sensitive to misallocations across states as we can observe dotted lines are closer to each other compared to the case with misallocations, presumably because of the firm ability to adapt across states thanks to its lower α . A similar effect can be observed on output, capital, investment, and dividends, where an increase in the firm's ability to choose technologies to adapt to exogenous shocks results in a more efficient allocation of resources.

Figure 1.9
Comparative Statistics Misallocation vs Standard Model



This graph shows the relevant dependent variables versus misallocation τ . $\tau > 0$ represents distortions as taxes, and $\tau < 0$ represents distortions as subsidies. High states are in blue and low-state in red. The model is simulated using the curvature parameter of $\alpha = 3.80$ (solid lines) and $\alpha = 1.66$ (dotted lines). Each point corresponds to the mean realization of: (top-left) firm value; (top-right) output; (middle-left) capital; (middle-right) investment; (bottom-left) dividend; (bottom-right) expected investment return.

1.6 Conclusion

I analyze firms' misallocation through the output distortions channel, using the Belo (2010) model as a framework. I calibrate the curvature parameter α and the production factor loading λ using a two-stage GMM estimation. The model is overidentified and the J -test is not rejected. I calibrate the model using standard Fama French test portfolios and find in the cross-section the estimated curvature parameter is larger compared to the original values obtained in Belo (2010). This implies misallocations

reduce the firm's ability to respond to the different states of nature. I calibrate and solve the model in the special case of a single representative firm. In particular, the model's solution sheds light on the effect of misallocations on relative production allocations, investment, and value. The impact of misallocation on the representative firm is larger when misallocations are present in the low-state (recessions) compared to high-states (expansions). Finally, I simulate the model in two cases: with the estimated curvature parameter considering misallocation ($\alpha = 3.80$) and using the standard Belo model ($\alpha = 1.66$). I find that the impact of misallocation on firm value, production, capital, investment, and investment return is larger when for $\alpha = 3.80$. This indicates that firms may be less agile to adapt across states of nature and provides more evidence of the detrimental effect of misallocations.

APPENDICES

1.A Proofs and Derivation

1.A.1 Producer's Maximization Problem

The firm solves

$$V(\mathbf{X}_{jt}) = \max_{I_{jt}, \varepsilon_{jt+1}} \{D_{jt} + \mathbb{E}_t[M_{t+1}V(\mathbf{X}_{jt+1})]\} \quad (1.46)$$

where $\mathbf{X}_{jt} \equiv (K_{jt}, \varepsilon_{jt}, P_{jt}, \mathbf{Z}_{jt})$ and \mathbf{Z}_{jt} contains all forecasting variables. The firm is subject to the following constraints:

$$D_{jt} = P_{jt}Y_{jt}(1 - \tau_{jt}) - I_{jt} \quad (1.47)$$

$$K_{jt+1} = I_{jt} + (1 - \delta)K_{jt} \quad (1.48)$$

$$1 \geq \mathbb{E}_t \left[\left(\frac{\varepsilon_{jt+1}}{\theta_{jt+1}} \right)^\alpha \right]^{\frac{1}{\alpha}} \quad (1.49)$$

$$Y_{jt+1} = \varepsilon_{jt+1} F^j(K_{jt+1}) \quad (1.50)$$

First-order conditions:

$$1 = \mathbb{E}_t [M_{t+1} V_{K_j}(\mathbf{X}_{jt+1})] \quad (1.51)$$

$$M_{t+1} V_{\varepsilon_j}(\mathbf{X}_{jt+1}) = \mu_{jt} \left(\mathbb{E}_t \left[\left(\frac{\varepsilon_{jt+1}}{\theta_{jt+1}} \right)^\alpha \right] \right)^{\frac{1-\alpha}{\alpha}} \varepsilon_{jt+1}^{\alpha-1} \theta_{jt+1}^{-\alpha} \quad (1.52)$$

Firms always maximize production, so the constraint (1.49) is binding. This gives:

$$M_{t+1}V_{\varepsilon_j}(\mathbf{X}_{jt+1}) = \mu_{jt}\varepsilon_{jt+1}^{\alpha-1}\theta_{jt+1}^{-\alpha} \quad (1.53)$$

Taking envelope conditions:

$$V_{K_j}(\mathbf{X}_{jt}) = P_{jt}\varepsilon_{jt}F_{K_j}^j(K_{jt})(1 - \tau_{jt}) + \mathbb{E}_t[M_{t+1}V_{K_j}(\mathbf{X}_{jt+1})](1 - \delta) \quad (1.54)$$

$$V_{\varepsilon_{jt}}(\mathbf{X}_{jt}) = P_{jt}F^j(K_{jt})(1 - \tau_{jt}) \quad (1.55)$$

Replacing the FOC (1.51) into the envelope condition (1.54) at $t + 1$

$$V_{K_j}(\mathbf{X}_{jt+1}) = P_{jt+1}\varepsilon_{jt+1}F_{K_j}^j(K_{jt+1})(1 - \tau_{jt+1}) + (1 - \delta) \quad (1.56)$$

Defining the investment return as

$$R_{jt+1}^I = V_{K_j}(\mathbf{X}_{jt+1}) \quad (1.57)$$

Replacing back into (1.51)

$$\mathbb{E}_t[M_{t+1}R_{jt+1}^I] = 1 \quad (1.58)$$

Replacing (1.55) at time $t + 1$ into (1.53)

$$M_{t+1}P_{jt+1}F^j(K_{jt+1})(1 - \tau_{jt+1}) = \mu_{jt}\varepsilon_{jt+1}^{\alpha-1}\theta_{jt+1}^{-\alpha}. \quad (1.59)$$

To derive the firm's choices under different state contingent misallocations, first, plug the envelope condition from Eq. (1.55) into the simplified first order condition

denoted as Eq. (1.53) to derive the Lagrange multiplier μ_{jt}

$$\mu_{jt} = \frac{\mathbb{E}_t [M_{t+1} P_{jt+1} F^j(K_{jt+1}) (1 - \tau_{jt+1})]}{\mathbb{E}_t [\varepsilon_{jt+1}^{\alpha-1} \theta_{jt+1}^{-\alpha}]} \quad (1.60)$$

Thus, the SDF from the firms' optimal choice of productivity level is:

$$M_{t+1} = \eta_{jt} \frac{1}{(1 - \tau_{jt+1})} \left(\frac{P_{jt+1}}{P_{jt}} \right)^{-1} \left(\frac{\varepsilon_{jt+1}}{\varepsilon_{jt}} \right)^{\alpha-1} \left(\frac{\theta_{jt+1}}{\theta_{jt}} \right)^{-\alpha}, \quad (1.61)$$

where

$$\eta_{jt} = \frac{\mathbb{E}_t \left[M_{t+1} \left(\frac{P_{jt+1}}{P_{jt}} \right) (1 - \tau_{jt+1}) \right]}{\mathbb{E}_t \left[\left(\frac{\varepsilon_{jt+1}}{\varepsilon_{jt}} \right)^{\alpha-1} \left(\frac{\theta_{jt+1}}{\theta_{jt}} \right)^{-\alpha} \right]}. \quad (1.62)$$

I can also obtain explicitly the value of η_{jt} replacing the SDF from Eq. (1.61) in Eq. (1.58)

$$\eta_{jt} = \left(\mathbb{E} \left[\frac{1}{(1 - \tau_{jt+1})} \left(\frac{P_{jt+1}}{P_{jt}} \right)^{-1} \left(\frac{\varepsilon_{jt+1}}{\varepsilon_{jt}} \right)^{\alpha-1} \left(\frac{\theta_{jt+1}}{\theta_{jt}} \right)^{-\alpha} R_{jt+1}^I \right] \right)^{-1} \quad (1.63)$$

References

- Belo, Frederico. 2007. “A pure production-based asset pricing model.” *Working Paper* (August): 1–47.
- . 2010. “Production-based measures of risk for asset pricing.” *Journal of Monetary Economics* 57 (2): 146–163. ISSN: 03043932. <https://doi.org/10.1016/j.jmoneco.2009.12.001>.
- Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry. 2018. “Really Uncertain Business Cycles.” *Econometrica* 86 (3): 1031–1065. <https://doi.org/https://doi.org/10.3982/ECTA10927>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA10927>.
- Campbell, John Y. 2003. “Chapter 13 Consumption-based asset pricing.” In *Financial Markets and Asset Pricing*, 1:803–887. Handbook of the Economics of Finance. Elsevier. [https://doi.org/https://doi.org/10.1016/S1574-0102\(03\)01022-7](https://doi.org/https://doi.org/10.1016/S1574-0102(03)01022-7).
- Campbell, John Y. 2017. *Financial decisions and markets: a course in asset pricing*. Princeton University Press.
- Cochrane, John H. 1996. “A Cross-Sectional Test of an Investment-Based Asset Pricing Model.” *Journal of Political Economy* 104 (3): 572–621. ISSN: 0022-3808. <https://doi.org/10.1086/262034>.
- Cochrane, John H. 2009. *Asset pricing: Revised edition*. Princeton university press.

- Cochrane, John H. 1991. "Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations." *The Journal of Finance* 46 (1): 209–237. ISSN: 00221082. <https://doi.org/10.1111/j.1540-6261.1991.tb03750.x>.
- . 1988. *Production Based Asset Pricing*. NBER Working Papers 2776. National Bureau of Economic Research, Inc, November. <https://ideas.repec.org/p/nbr/nberwo/2776.html>.
- Cochrane, John H. 1993. "Rethinking production under uncertainty." *Working Paper*, 0–56. <http://faculty.chicagobooth.edu/john.Cochrane/research/papers/CochraneRethinkingProductionUnderUncertainty.pdf>.
- Cochrane, John H. 2021. "Rethinking Production under Uncertainty." *The Review of Asset Pricing Studies* 11, no. 1 (March): 1–59. ISSN: 2045-9920. <https://doi.org/10.1093/rapstu/raaa006>. eprint: https://academic.oup.com/raps/article-pdf/11/1/1/36267995/raaa006_supplementary_data.pdf.
- Duffie, Darrell. 2010. *Dynamic asset pricing theory*. Princeton University Press.
- Fernald, John. 2014. "A quarterly, utilization-adjusted series on total factor productivity." Federal Reserve Bank of San Francisco.
- Gomes, João F., Leonid Kogan, and Motohiro Yogo. 2009. "Durability of Output and Expected Stock Returns." *Journal of Political Economy* 117 (5): 941–986. <https://doi.org/10.1086/648882>. eprint: <https://doi.org/10.1086/648882>.
- Guo, Bin, Han Zhang, and Yongjie Zhang. 2022. "A production-based asset pricing model with R&D investment." *Available at SSRN*.

- Hall, Robert E. 1988. “Intertemporal substitution in consumption.” *Journal of political economy* 96 (2): 339–357.
- Hansen, Lars Peter, and Kenneth J. Singleton. 1982. “Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models.” *Econometrica* 50 (5): 1269–1286. ISSN: 00129682, 14680262, accessed April 6, 2023. <http://www.jstor.org/stable/1911873>.
- Hsieh, Chang-Tai, and Peter J. Klenow. 2009. “Misallocation and Manufacturing TFP in China and India*.” *The Quarterly Journal of Economics* 124, no. 4 (November): 1403–1448. ISSN: 0033-5533. <https://doi.org/10.1162/qjec.2009.124.4.1403>. eprint: <https://academic.oup.com/qje/article-pdf/124/4/1403/5407260/124-4-1403.pdf>.
- Ljungqvist, Lars, and Thomas J. Sargent. 2012. *Recursive Macroeconomic Theory, Third Edition*. Vol. 1. MIT Press Books 0262018748. The MIT Press. ISBN: ARRAY(0x55446ac8). <https://ideas.repec.org/b/mtp/titles/0262018748.html>.
- Restuccia, Diego, and Richard Rogerson. 2008. “Policy distortions and aggregate productivity with heterogeneous establishments.” *Review of Economic Dynamics* 11 (4): 707–720. ISSN: 1094-2025. <https://doi.org/https://doi.org/10.1016/j.red.2008.05.002>.

CHAPTER 2

Earnings Expectations and Asset Prices

With Denis Mukanov (*UCLA Anderson School*) and **Danyu Zhang** (*UCLA Anderson School*)

2.1 Introduction

Bouchaud et al. (2019) use the methodology developed by Coibion and Gorodnichenko (2015) to examine the rationality of financial analysts' earnings forecasts and find that analysts' forecasts underreact to news unconditionally. An additional finding in the literature, most notably in Coibion and Gorodnichenko (2015), is that there is evidence for state dependence in the expectation formation process. In this paper, we examine the degree to which the state dependence of information rigidity extends to equity analysts' earnings expectations. We use financial analysts' earnings per share (EPS) forecasts from the Thomson Reuters Institutional Brokers Estimate System (I/B/E/S) and document the following stylized facts: (1) unconditionally, expectations underreact to news, in line with the findings of Abarbanell and Bernard (1992) and Bouchaud et al. (2019); (2) the stickiness of earnings expectations de-

clines significantly during periods of high market volatility; and (3) the stickiness of earnings expectations declines significantly over our sample period (01/1986 to 12/2021).¹

We develop a simple model featuring rational inattention, as in Mankiw and Reis (2002), to explain the state dependence in equity analysts' expectation formation process. A fraction of the market participants choose to optimally remain inattentive to private and public signals. The presence of inattentive agents causes the average (consensus) earnings expectations to underreact to news. In our framework, the time-varying costs and benefits of being attentive drive the state-dependence of expectation stickiness. The benefits of acquiring information increase during high-volatility periods, which translates into lower expectation stickiness during high-volatility periods. Our model is also able to account for expectation stickiness declining over our sample, given recent technological innovations that have likely reduced the cost of information acquisition.

Given the ability of our model to replicate the conditional behavior of equity analysts' earnings expectations, we focus on its asset pricing implications. The momentum anomaly (Jegadeesh and Titman 1993) provides a natural starting point for our analysis. Momentum is one of the most robust empirical results in financial economics. Consequently, it has garnered significant attention in the asset pricing literature. Our framework fits into the class of models that relate the profitability of momentum to the slow diffusion of fundamental information (Hong and Stein 1999): the presence of inattentive agents causes prices not to incorporate all publicly available information immediately. Instead, stock prices underreact news, so that prices

1. Throughout this paper, we use the terms information rigidity and expectation stickiness interchangeably as the two terms are equivalent within our theoretical framework.

are, on average, too low following positive news and too high following negative news. As a consequence, positive lagged returns predict high subsequent returns and negative lagged returns predict low subsequent returns, thus generating return patterns consistent with the unconditional profitability of momentum.

More recently, Daniel and Moskowitz (2016) find that momentum experiences negative returns (crashes) following high-volatility periods and Chordia, Subrahmanyam, and Tong (2014) find that the profitability of momentum has diminished significantly over the period between 1976 and 2011. First, we verify the robustness of these results in our extended sample. Then, we show that our model is capable of accounting for these results. According to our model, the profitability of momentum is tied to the stickiness of market participants' expectations. The increasing relative cost of being inattentive during high-volatility periods accounts for momentum crashes and the declining cost of information acquisition accounts for the attenuating profitability of momentum over our sample.

A prediction of our model is that the relative profitability of momentum strategies with different lookback periods differs based on the level of market volatility. During low-volatility periods, momentum strategies with longer lookback periods (e.g. the Jegadeesh and Titman 1993 $t - 12$ to $t - 2$ strategy) tend to outperform momentum strategies with shorter lookback periods (e.g. a $t - 3$ to $t - 2$ momentum strategy). However, during high-volatility episodes, short-run momentum strategies may deliver higher returns than long-run momentum strategies. Based on this idea, we propose a trading strategy that mixes long-run and short-run momentum strategies (similar to Goulding, Harvey, and Mazzoleni 2022), with greater weight placed on the short-run strategy during high-volatility periods. The resultant mixed momentum strategy lessens the impact of momentum crashes and earns a significant α with respect to

the 12-2 momentum strategy.

A prediction our model shares with any model featuring deviations from full-information rational expectations is that the wedge between the objective earnings expectations and analysts' forecasts predicts stock returns. In order to test this model prediction, we use the Extreme Gradient Boosting machine learning algorithm to construct a proxy for objective expectations (similar to Van Binsbergen, Han, and Lopez-Lira 2022). Our approach serves as an extension of papers in the literature that use linear regression frameworks to extract the predictable component of analysts' forecast errors (e.g. So 2013 and Frankel and Lee 1998).

Armed with a measure of objective expectations, we sort stocks into portfolios based on the value of the predictable component of forecast errors. We call the long-short portfolio that is long on the stocks with the most pessimistic earnings expectations and short on the stocks with the most optimistic earnings expectations pessimistic-minus-optimistic (PMO). The PMO strategy generates an annualized Sharpe ratio of 1.16 and its returns cannot be fully explained by standard multifactor models.

Related literature

Our paper contributes to a number of different strands of the literature. In terms of documenting state-dependent expectation stickiness, our paper is most closely related to Coibion and Gorodnichenko (2015). In other related papers, Loungani, Stekler, and Tamirisa (2013) show that professional forecasters increase the rate at which they incorporate news into their forecasts as the economy enters a recession

and Andrade and Le Bihan (2013) document that the Great Recession is associated with increased attentiveness to unemployment, real GDP, and inflation news among professional forecasters surveyed by the European Central Bank.

Our work is also related to the literature that documents systematic errors in equity analysts' earnings expectations and relates the systematic errors to the profitability of various trading strategies. For instance, Bouchaud et al. (2019) show that analysts' short-run earnings expectations underreact to news and propose an explanation for the profitability anomaly (Novy-Marx 2013) based on underreaction to earnings surprises. On the other hand, Bordalo et al. (2022) show that analysts' long-term earnings growth expectations overreact to news and develop a model in which the profitability of the Fama and French (2015) factors is driven by the overreaction of long-term expectations. Engelberg, McLean, and Pontiff (2018) use *ex-post* forecast errors and show that analysts tend to have overly optimistic (pessimistic) expectations for stocks in the short (long) leg of anomalies.

Our framework is based on the idea that costly information acquisition causes earnings shocks to be incorporated into consensus expectations slowly. The slow diffusion of information, in turn, causes momentum. It is well-established in the asset pricing literature that limited investor attention is associated with slow diffusion of information and underreaction to news. For instance, Ben-Rephael, Da, and Israelsen (2017) provide a measure of abnormal institutional attention and show that the post-earnings announcement drift is driven by announcements which do not receive sufficient attention from institutional investors. In related work, Hirshleifer, Lim, and Teoh (2009) show that underreaction to earnings announcements is stronger on days with a greater number of earnings announcements and Dellavigna and Pollet (2009) show that underreaction is stronger for earnings announcements that take

place on Friday. Hong, Lim, and Stein (2000) and Da, Guren, and Warachka (2014) propose two distinct proxies for the speed at which news is incorporated into consensus expectations, residual analyst coverage and information discreteness, respectively. Both papers find that stocks for which information is incorporated into expectations more slowly are associated with higher momentum returns.

In terms of the use of non-linear methods to construct *ex-ante* forecast errors, our paper is most closely related to Van Binsbergen, Han, and Lopez-Lira (2022), Silva and Thesmar (2023), and Cao and You (2021). Van Binsbergen, Han, and Lopez-Lira (2022) find that stocks with upward- (downward-) biased earnings forecasts tend to earn lower (higher) returns going forward. Silva and Thesmar (2023) decompose analysts' forecast errors at different horizons into soft information, forecast bias, and forecast noise. Cao and You (2021) document that earnings information uncovered by machine learning algorithms (over extant models) is significantly associated with future stock returns and earnings forecast errors.

2.2 Data

2.2.1 Analysts' forecasts

We obtain consensus (median) earnings-per-share (EPS) forecasts from the I/B/E/S Unadjusted Summary file. Following Bouchaud et al. (2019), we focus on the one-year and two-year earnings forecasts.² I/B/E/S updates earnings forecasts monthly. Our tests of forecast error predictability are based on the forecasts immediately following

2. The forecasting horizon is identified using the I/B/E/S Forecast Period Indicator variable *FPI*.

the announcement of the previous fiscal year’s earnings.

We match earnings forecasts with earnings realizations from the I/B/E/S actual file using ticker and fiscal end date.³ Before merging the two datasets, we adjust the realized EPS values for stock splits using the CRSP cumulative adjustment factor CFACSHR (Diether, Malloy, and Scherbina 2002):

$$AdjustedEPS_{f,t} = \frac{CFACSHR_{f,t-1}}{CFACSHR_{f,t}} \times EPS_{f,t}$$

The forecast error predictability tests are based on firms with fiscal year ends between 01/1986 and 12/2021. Our final dataset contains 78,287 firm-year observations.

2.2.2 Stock and trading strategy returns

We obtain monthly return and stock price data from CRSP. We start with all firms in the monthly CRSP database between 1986 and 2021 and apply the following filters: we only keep the common stocks (share codes 10 and 11) of firms listed on the NYSE, Amex, or Nasdaq (exchange code 1, 2, and 3). We also exclude firms whose stock price is below \$1. We then match the CRSP data with the analyst forecast data described in the previous section.⁴

In this paper, we also utilize a number of off-the-shelf trading strategies and factor returns, which serve as control variables or building blocks for the strategies

3. Fiscal end dates are denoted by PENDS in the actual file and by FPEDATS in the summary file.

4. We merge I/B/E/S data with CRSP data using the link table provided by Wharton Research Data Services.

proposed in this paper. We make the decision to use data used in previous research to ensure greater comparability with existing work. We obtain the returns of the trading strategies from one of two sources: Kenneth French’s data library or the Global Factor Data repository (Jensen, Kelly, and Pedersen 2022).

2.3 Forecast Error Predictability

We start our analysis by examining the ability of earnings forecast revisions to predict earnings forecast errors. The resultant regression coefficient allows us to draw conclusions regarding the degree of information rigidity in equity analysts’ expectations (Coibion and Gorodnichenko 2015).

To examine the ability of forecast revisions to predict forecast errors, we first construct forecast revisions. Forecast revisions are defined as the difference between the time t consensus forecast for firm f ’s fiscal year τ earnings (one-year forecast) and the time $t - 1$ consensus forecast for firm f ’s fiscal year τ earnings (two-year forecast). Throughout our analysis, we use τ to denote fiscal years and t to denote calendar time.

Following Bouchaud et al. (2019), we normalize the revision by firm f ’s stock price in year $t - 1$, $P_{f,t-1}$.⁵ Therefore,

$$FR_{f,t} = \frac{\mathbb{F}_t[e_{f,\tau}] - \mathbb{F}_{t-1}[e_{f,\tau}]}{P_{f,t-1}}$$

where e denotes earnings per share.

5. $P_{f,t-1}$ is the stock price at the end of the month used to determine $\mathbb{F}_{t-1}[e_{f,\tau}]$.

Forecast errors are defined as the difference between the actual fiscal year τ earnings and the year t earnings forecast, normalized by the time $t - 1$ stock price:

$$FE_{f,\tau} = \frac{e_{f,\tau} - \mathbb{F}_t[e_{f,\tau}]}{P_{f,t-1}}$$

We begin our analysis by estimating the following regression, following Bouchaud et al. (2019):

$$FE_{f,\tau} = \alpha^{CG} + \beta^{CG} FR_{f,t} + \varepsilon_{t,\tau} \quad (2.1)$$

where we winsorize the forecast errors and forecast revisions at the 1% and 99% levels.

If the analysts' information set includes all information available at time t , we would not be able to predict forecast errors using any time t information, including forecast revisions, i.e. the regression coefficient, β^{CG} , is equal to zero under the null of rational expectations. If analysts' expectations underreact to earnings shocks, we would observe $\hat{\beta}^{CG} > 0$. The mechanism underlying this result is the following: let us assume that agents receive a piece of positive news at time t . This implies that $FR_{f,t} > 0$. However, if agents underreact to the news, the forecast errors will also be positive, on average, i.e. forecast revisions will be positively correlated with forecast errors.

Conversely, overreaction to news implies $\hat{\beta}^{CG} < 0$. The mechanism underlying this result is the same as the one outlined in the underreaction case. The difference is that the forecast errors will be negative, on average, following positive news. That is, forecast errors are negatively correlated with forecast revisions in the case of overreaction.

The results in the main body of this paper are based on a panel regression of all firm-level observations. The results of the baseline regression are reported in the first column of Table 2.1. We estimate the β^{CG} coefficient to be positive ($\hat{\beta}^{CG} = 0.167$) and highly statistically significant (t -statistic = 5.195). The positive regression coefficient indicates that equity analysts' expectations underreact to earnings shocks. This result is consistent with the findings in Bouchaud et al. (2019) who estimate β^{CG} to be 0.165.

Volatility and information rigidity

Coibion and Gorodnichenko (2015) show that the rigidity of SPF survey respondents' expectations declines during NBER recessions. They hypothesize that the decline in information rigidity is driven by the high macroeconomic volatility prevailing during recessions. In order to test the validity of this hypothesis in the context of earnings expectations we examine the relation between information rigidity and stock market volatility. In particular, we estimate the following regression:

$$FE_{f,\tau} = \alpha^{CG} + \alpha_{\Delta}^{CG} z_t + (\beta^{CG} + \beta_{\Delta}^{CG} z_t) FR_{f,t} + \varepsilon_{t,\tau} \quad (2.2)$$

where z_t is either time t stock market volatility ($\hat{\sigma}_{\text{mkt}}$) or an indicator, which takes on the value of 1 if stock market volatility at time t is above a certain threshold ($\mathbf{1}_{\text{HV}}$). In our analysis, z_t is computed by taking the average of beginning-of-month and end-of-month volatility for month t . Alternative methods of computing volatility produce qualitatively similar results.

The results from the estimation of the regression in Equation (2.2) are reported

in columns (2) through (7) of Table 2.1. In this table, we consider two measures of volatility: realized volatility (RV) and implied volatility (IV). RV is the standard deviation of the daily market returns over the 126 days prior to time t (Daniel and Moskowitz 2016):

$$\hat{\sigma}_{mkt,t} = \sqrt{\sum_{d=1/126}^1 \left(R_{mkt,t+d} - \frac{\sum_{d=1/126}^1 R_{mkt,t+d}}{126} \right)^2} \quad (2.3)$$

where $R_{mkt,t}$ represents the day t return of value-weighted CRSP.

We use the VXO (implied volatility of S&P 100 index options) as a measure of IV. The "high volatility" indicators ($\mathbf{1}_{HV}$) considered in the table take on the value of 1 if volatility is above the 80th or 90th percentile of the volatility estimates within our full sample.

The results in Table 2.1, show that information rigidity declines significantly during high-volatility periods: we estimate β_{Δ}^{CG} coefficients to be negative and statistically significant. The two volatility measures considered in the table generate similar results. We estimate $\hat{\beta}_{\Delta}^{CG}$ to be -0.056 using implied volatility. The effect of using realized volatility is slightly weaker compared to implied volatility.

Information rigidity over time

Coibion and Gorodnichenko (2015) document a low-frequency variation in the stickiness of macroeconomic expectations. In particular, they find that the rigidity of SPF respondents' expectations increased significantly during the period characterized by low macroeconomic volatility known as the Great Moderation.

In this section, we examine if equity analysts' earnings expectations display similar low-frequency patterns. To do so, we estimate the regression in (2.1) using overlapping two-year windows.⁶ Figure 2.1 plots the dynamics of the β_t^{CG} coefficients obtained using this methodology as well as the associated 95% confidence intervals and a fitted trend line. Figure 2.1 indicates that the rigidity of earnings expectations declines significantly throughout the sample. In fact, we cannot reject rational expectations during the post-2010 sample ($\hat{\beta}^{CG} = 0.077$, t -statistic = 1.616).

The lack of stickiness in analysts' earnings expectations during the latter parts of our sample is consistent with the findings of Martineau (2023) who shows that the post-earnings announcement drift has disappeared in 2006 for large stocks and in the 2010s for microcaps.

2.4 Model

In Section 2.3, we identify the following patterns in equity analysts' consensus earnings expectations:

1. Equity analysts' earnings expectations underreact to shocks unconditionally.
2. The degree of underreaction to earnings shocks declines significantly during high volatility periods.
3. The degree of underreaction to earnings shocks declines over our sample.

6. For instance, the $\hat{\beta}^{CG}$ coefficient associated with 1986 is based on a sample that includes 1986 and 1987.

In this section, we develop a simple model capable of accounting for the three stylized facts presented above.

The economy we study is based on the one in Pouget, Sauvagnat, and Villeneuve (2016). It consists of two assets: a riskless asset in perfectly elastic supply and a risky asset in fixed supply, x . The gross return of the riskless asset is normalized to 1. There are $T - 1$ dates of trading, indexed by $t \in \{0, 1, \dots, T - 1\}$. Consumption takes place at date T during which the payoff of the risky asset is drawn from a normal distribution with a mean of 0 and variance of σ_V^2 . There is no consumption prior to the final period and there is no time discounting.

We set $T = 3$, which allows us to derive closed-form solutions for moments of interest. The signal structure we consider follows Daniel, Hirshleifer, and Subrahmanyam (1998). At time $t = 1$, overconfident agent i (in measure one) obtains a private signal $S_{i,1}$ with probability λ_1 that is determined endogenously. Fraction $1 - \lambda_1$ of the market participants fail to obtain the private signal. The inattentive agents use their prior beliefs to form expectations about the future, as in Mankiw and Reis (2002). Following the standard approach in the literature, we assume that $S_{i,1} = V + \varepsilon_{i,1}$. Where the noise term, $\varepsilon_{i,1}$, is independently distributed across agents and normally distributed with a mean of 0 and a variance of σ_S^2 . Following Daniel, Hirshleifer, and Subrahmanyam (1998), we model overconfidence as agents overassessing the quality of their private signals, i.e. agents perceive the variance of their private signal to be $\sigma_C^2 < \sigma_S^2$.⁷

At time $t = 2$, a public signal S_2 is realized. Agent i observes the signal and

7. Overconfidence has limited bearing on the ability of our model to replicate the patterns in equity analysts' earnings expectations. Overconfidence plays an important role in Section 2.5, in which we study the asset pricing implication of our model.

updates her expectations with probability $\lambda_{i,2}$, which depends on the agent's information set at the beginning of period 2. The public signal is equal to $V + \varepsilon_2$ where $\varepsilon_2 \sim N(0, \sigma_S^2)$. We assume that agents perceive the precision of the public signal correctly.

Figure 2.2 provides a graphical illustration of our model. To relate our model to the empirical results in Section 2.3, we interpret the private signals as a representation of the information collected by equity analysts prior to an earnings release. The public signal and the final asset payoff represent earnings announcements.

In order to focus on the expectation formation process presented in this paper, we make several assumptions that follow Pouget, Sauvagnat, and Villeneuve (2016). First, we assume that the market participants are risk-neutral and incur an exogenous trading cost that is quadratic in their portfolio positions, i.e. the total cost for trader i is $\frac{\psi}{2} q_{i,t}^2$. We also assume that agents cannot costlessly extract signals from market prices.⁸ Finally, we assume that market participants agree to disagree and trade at the prevailing market price during periods 1 and 2.

2.4.1 Portfolio choice problem

Investors in our model derive utility from end-of-life consumption, C_3 . The end-of-life consumption of agent i can be represented as $\sum_{t=0}^3 q_{i,t} (V - p_t)$.

Therefore, agent i 's portfolio choice problem can be written as:

8. The framework we have in mind involves inattentive agents submitting demand schedules to a Walrasian auctioneer. These demand schedules are only revised during periods during which agents acquire a new signal.

$$\max_{q_{i,t}} \left\{ \sum_{t=1}^3 U_{i,t} \right\} \equiv \max_{q_{i,t}} \left\{ \sum_{t=1}^3 \left(q_{i,t} \mathbb{E}_t^i [V - p_t] - \frac{\psi}{2} q_{i,t}^2 \right) \right\} \quad (2.4)$$

As shown by Pouget, Sauvagnat, and Villeneuve (2016), this maximization problem is equivalent to maximizing utility for each period separately.

2.4.2 Information acquisition problem

At the beginning of period 1, investor i chooses the probability λ_1 with which she acquires signal $S_{i,1}$ and updates her expectations. At the beginning of period 2, agent i chooses the probability with which she acquires the public signal S_2 .

The cost of information acquisition is specified by the function $C(\lambda_{i,t})$. We assume that $C(\cdot)$ is a strictly increasing convex function of $\lambda_{i,t}$. In particular, we consider the following functional form:

$$C(\lambda_{i,t}) = \frac{\varphi}{\kappa + 1} \lambda_{i,t}^{\kappa+1}, \text{ with } \kappa > 1 \text{ and } \varphi > 0 \quad (2.5)$$

In this specification, the φ parameter shifts the marginal cost of information acquisition and κ influences the local curvature of the $C(\cdot)$ function.

Given the setup presented in this section, agents i 's choice of λ at the beginning of period t solves the following problem:

$$\max_{\lambda_{i,t}} \left\{ \mathbb{E}_{t-1}^i [U_{i,t}] - C(\lambda_{i,t}) \right\} \text{ for } t \in \{1, 2\} \quad (2.6)$$

subject to the constraint that $\lambda_{i,t} \in [0, 1]$.

2.4.3 Equilibrium

In order to solve the model, we follow the approach in Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) that involves three steps.

Step 1: Solve for optimal portfolios, given information sets.

Agent i 's time t demand takes on the following form:

$$\begin{aligned} q_{i,0} &= 0 \\ q_{i,t} &= \frac{\mathbb{E}_t^i[V] - p_t}{\psi} \text{ for } t = \{1, 2\} \end{aligned} \tag{2.7}$$

Step 2: Clear the asset market.

The market-clearing condition for the risky asset is:

$$\int_0^1 q_{i,t} di = x \tag{2.8}$$

Plugging in the expression for agent i 's demand into the market clearing condition, we obtain the following expression for the price of the risky asset:

$$p_t = \int_0^1 \mathbb{E}_t^i[V] di - \psi x = \bar{\mathbb{E}}_t[V] - \psi x \tag{2.9}$$

where we use the $\bar{\mathbb{E}}_t[\cdot]$ notation to denote average (consensus) expectations.

The market participants are homogeneous at the beginning of period 1. Then the period 1 price of the risky asset is:

$$p_1 = \lambda_1 \mathbb{E}_1[V] + (1 - \lambda_1) \mathbb{E}_0[V] - \psi x = \lambda_1 \mathbb{E}_1[V] - \psi x \tag{2.10}$$

where $\mathbb{E}_1[V] = \frac{\sigma_V^2}{\sigma_V^2 + \sigma_C^2} S_1$ represents the expectation of the attentive agents.

At the beginning of period 2, attentive (A) and inattentive (I) agents have different information sets. Then the time 2 price of the risky asset is::

$$p_2 = \lambda_1 \lambda_{2|A} \mathbb{E}_2^C[V] + (1 - \lambda_1) \lambda_{2|I} \mathbb{E}_2^P[V] + (1 - \lambda_{2|A}) \lambda_1 \mathbb{E}_1[V] + (1 - \lambda_1)(1 - \lambda_{2|I}) \mathbb{E}_0[V] - \psi x \quad (2.11)$$

where $\lambda_{2|A}$ and $\lambda_{2|I}$ represent the optimal updating probabilities for agents who are attentive and inattentive to the private signals, respectively. Here $\mathbb{E}_2^C[V] = \frac{\sigma_V^2 \sigma_S^2}{\sigma_V^2 \sigma_S^2 + \sigma_V^2 \sigma_C^2 + \sigma_C^2 \sigma_S^2} S_1 + \frac{\sigma_V^2 \sigma_C^2}{\sigma_V^2 \sigma_S^2 + \sigma_V^2 \sigma_C^2 + \sigma_C^2 \sigma_S^2} S_2$ and $\mathbb{E}_2^P[V] = \frac{\sigma_V^2}{\sigma_V^2 + \sigma_S^2} S_2$.⁹

Step 3: Solve for information choices.

In Appendix 2.A.1, we show that agent i 's optimal $\lambda_{i,t}$ for $t \in \{1, 2\}$ is the solution to the following equation:

$$\frac{\lambda_{i,t}}{\psi} \text{var}_{i,t-1}(\mathbb{E}_t[V]) - \varphi \lambda_{i,t}^\kappa = 0 \quad (2.12)$$

Based on the expression in (2.12), the optimal values of λ_1 , $\lambda_{2|A}$, and $\lambda_{2|I}$ solve the

9. An assumption underlying this expression is that agents who are inattentive in period 1 do not gain access to a private signal when they update their expectations in period 2.

following equations:

$$\begin{aligned}
\lambda_1: \quad & \frac{\sigma_V^4 (\sigma_V^2 + \sigma_S^2)}{\psi (\sigma_V^2 + \sigma_C^2)^2} \lambda - \varphi \lambda^\kappa = 0 \\
\lambda_{2|I}: \quad & \frac{\sigma_V^4}{\psi (\sigma_V^2 + \sigma_S^2)} \lambda - \varphi \lambda^\kappa = 0 \\
\lambda_{2|A}: \quad & \frac{\sigma_V^4 (\sigma_S^4 + \sigma_C^4) (\sigma_V^2 + \sigma_S^2)}{\psi (\sigma_V^2 (\sigma_C^2 + \sigma_S^2) + \sigma_C^2 \sigma_S^2)^2} \lambda - \varphi \lambda^\kappa = 0
\end{aligned} \tag{2.13}$$

2.4.4 Model implications for forecast error predictability

In the context of our model, the tests presented in Section 2.3 examine the reaction of consensus expectations to the public signal, S_2 . Therefore, the model-implied version of the forecast error predictability coefficients is:

$$\beta^{CG} = \frac{\text{cov} \left(V - \bar{\mathbb{E}}_2 [V], \bar{\mathbb{E}}_2 [V] - \bar{\mathbb{E}}_1 [V] \right)}{\text{var} \left(\bar{\mathbb{E}}_2 [V] - \bar{\mathbb{E}}_1 [V] \right)} \tag{2.14}$$

where $\bar{\mathbb{E}}_1$ and $\bar{\mathbb{E}}_2$ are the consensus expectations of V at time $t = 1$ and $t = 2$. The theoretical expectations correspond $\mathbb{F}_t [e_{f,\tau}]$ and $\mathbb{F}_{t-1} [e_{f,\tau}]$ from Section 2.3.

An analytical expression for the error predictability coefficient and its derivation are outlined in Appendix 2.A.2. The actual expression for β^{CG} is not particularly intuitive. Therefore, in this section, we focus on a model with a single trading period. In a one-period setting with a single public signal,

$$\beta^{CG} = \frac{1 - \lambda}{\lambda} \tag{2.15}$$

Consequently, the partial derivative of the forecast error predictability coefficient with respect to fundamental volatility takes on the following form:

$$\frac{\partial \beta^{CG}}{\partial \sigma_V^2} = -\frac{\frac{\partial \lambda}{\partial \sigma_V^2}}{\lambda^2} < 0 \quad (2.16)$$

The expression in Equation (2.16) shows that our model generates information rigidity patterns consistent with those documented in the previous section: our model predicts that high-volatility periods are associated with lower information rigidity. The logic underlying this relation extends to the multi-period version of our model.

Volatility and information rigidity

In the context of the stylized model developed in this paper, we study the effects of volatility on the information acquisition decision by examining the relation between the probability of signal acquisition in periods 1 and 2 and fundamental variance σ_V^2 .

In order to provide analytical expressions for λ , we focus on the non-zero solutions of the equations in (2.12) for the special case of $\kappa = 2$.¹⁰ The optimal probabilities of information acquisition take on the following form:

$$\begin{aligned} \lambda_1 &= \frac{\sigma_V^4 (\sigma_V^2 + \sigma_S^2)}{\psi \varphi (\sigma_V^2 + \sigma_C^2)^2} \\ \lambda_{2|I} &= \frac{\sigma_V^4}{\psi \varphi (\sigma_V^2 + \sigma_S^2)} \\ \lambda_{2|A} &= \frac{\sigma_V^4 (\sigma_S^4 + \sigma_C^4) (\sigma_V^2 + \sigma_S^2)}{\psi \varphi (\sigma_V^2 (\sigma_C^2 + \sigma_S^2) + \sigma_C^2 \sigma_S^2)^2} \end{aligned} \quad (2.17)$$

The comparative statics with respect to σ_V^2 are detailed in Appendix 2.A.3. All of

10. Through numerical solutions we find that the logic of this special case is generalizable to the $\kappa > 2$ case.

the derivatives with respect to σ_V^2 are positive, i.e. agents are willing to expend more resources to acquire signals during high-volatility periods. Assuming that the state of the technology determines the cost of acquiring information and is, therefore, fixed in the short-run (i.e. it is unaffected by the state of the business cycle), our model produces results that are consistent with the patterns of information rigidity documented in Section 2.3.

Information rigidity over time

The assumption that the cost of information acquisition at time t is determined by the state of the technology at time t also allows the model to account for the decline in information rigidity between 1986 and 2021.

In the context of our framework, the φ parameter governs the cost of information acquisition. We model technological innovations as lower value of φ . Lower values of φ correspond to higher values of λ_1 , $\lambda_{2|A}$, and $\lambda_{2|I}$, as we indicated by the expressions in (2.17).

2.5 Asset Pricing Implications

Our model generates patterns of information rigidity that are consistent with the stylized facts documented in Section 2.3. In this section, we explore the asset pricing implications of the model.

2.5.1 Momentum

The model developed in the previous section provides us with three momentum signals, which we term long-run momentum signal $(p_2 - p_0)$ and short-run momentum signals $(p_1 - p_0)$ and $p_2 - p_1$.

Following Luo, Subrahmanyam, and Titman (2020), we define the unconditional short-run momentum parameter as:

$$\text{MOM}_S = \frac{\text{cov}(p_2 - p_1, p_1 - p_0) + \text{cov}(V - p_2, p_2 - p_1)}{2} \quad (2.18)$$

and the long-run momentum parameter as:

$$\text{MOM}_L = \text{cov}(V - p_2, p_2 - p_0) \quad (2.19)$$

We derive closed-form expressions for the momentum parameters in Appendix 2.A.4. In order to build intuition regarding the predictions of the model, we rely on simulations involving representative price paths. The average price path based on 10,000 simulated paths is shown in Figure 2.3. In this example, we focus on the case of a positive innovation, $V > 0$. The case with a negative innovation is entirely symmetric.

The figure is based on the following parameters: $\sigma_{V,low}^2 = 0.35$, $\sigma_{V,high}^2 = 0.6$, $\sigma_S^2 = 0.7$, $\sigma_C^2 = 0.3$. The transaction cost parameter ψ is set to 1 and the total supply of the risky asset, x , is set to 10^{-5} . We set the probability of a low volatility state is set to 0.90. The parameters associated with the cost function are $\varphi = 0.7$ and $\kappa = 4$. This combination of parameters is chosen to generate probabilities of

information acquisition roughly in line with those documented in Section 2.3.

As indicated in Figure 2.3, the model developed in Section 2.4 generates both short-run and long-run momentum. In the context of our model, the consensus reaction to news is determined by the interaction of two effects, which push expectations in opposite directions: the rational inattention effect generates underreaction to both private and public signals as information is incorporated into aggregate expectations with a delay (Hong and Stein 1999). On the other hand, the overconfidence effect generates overreaction to private signals, as in Daniel, Hirshleifer, and Subrahmanyam (1998). Given our parameter choices, the rational inattention effect dominates the overconfidence effect unconditionally and we observe underreaction in both the short run and the long run, i.e. both MOM_S and MOM_L are positive. The result is consistent with the fact that both slow (e.g. 12-2) and fast (e.g. 7-2) momentum strategies are profitable unconditionally.

2.5.2 Momentum and volatility

Daniel and Moskowitz (2016) show that long-run (12-month) momentum tends to have low returns following high volatility periods. In Table 2.2, we verify that these results continue to hold in our extended sample. In particular, we estimate the following regression to examine the effects of volatility on the profitability of momentum:

$$r_{WML,t} = \alpha + \alpha_{\Delta} z_{t-1} + \varepsilon_t \quad (2.20)$$

where $r_{WML,t}$ is the return of the 12-2 momentum strategy, obtained from Kenneth French's data library and z_{t-1} is one of the volatility-related variables considered in

Section 2.3 (realized/IMPLIED volatility or a high volatility dummy).

The results in Table 2.2 are consistent with the findings of Daniel and Moskowitz (2016): momentum earns significant positive returns in the order of 10.6% per year over our sample, unconditionally. The momentum strategy delivers significantly lower returns following high-volatility periods. Following periods which are in the top 20% in terms of volatility, momentum returns tend to be negative. This result is also consistent with the findings of Barroso and Wang (2022) who show that the momentum profits occur after periods of low volatility. In Appendix 2.B.1, we repeat the tests presented in this table using the $t - 12$ to $t - 1$ momentum factor from the Jensen-Kelly-Pedersen data repository. The results obtained using the Jensen-Kelly-Pedersen momentum factor mirror those in Table 2.2.

To understand the implications of our model for the impact of volatility on the profitability of momentum, we simulate our model separately for periods of high and low volatility. The results are presented in Figure 2.4. The parameters used to generate the figure are discussed in the previous subsection.

The mechanism that allows our model to generate diminishing momentum profitability during high-volatility episodes goes through the information acquisition channel. Higher values of σ_V^2 translate into higher probabilities of information acquisition, as shown in the equations in (2.17). As information is incorporated into the aggregate expectations faster, the opportunity for momentum profits diminishes.

If volatility is high enough, the overconfidence effect comes to dominate the rational inattention effect and the consensus expectations overreact to the private signals. The initial overreaction to the private signals is partially corrected during the subsequent trading period but p_2 does not revert all the way to its rational level, i.e. the

initial overreaction is further corrected during period 3.

Our model also provides a rationale for volatility management enhancing the profitability of momentum (Barroso and Santa-Clara 2015; Moreira and Muir 2017): the volatility-managed momentum strategy avoids the large losses of the baseline momentum strategy by limiting investors' exposure to momentum during high-volatility episodes and increasing investors' exposure to momentum during low-volatility episodes when momentum profits tend to be high.

2.5.2.1 Mixed momentum strategy

According to our model, the relative profitability of the long-run and short-run momentum strategies varies with fundamental volatility. During low-volatility periods, the long-run momentum strategy dominates the short-run momentum strategy due to the fact that the short-run momentum signal predicts profits with a low signal-to-noise ratio. The long-run momentum signal has a higher signal-to-noise ratio as the noise components of the $t = 1$ signal and the $t = 2$ signal cancel each other out on average.

On the other hand, the short-run momentum signal at time $t = 2$ is positively correlated with the price appreciation at time $t = 3$ ($V - p_2$).¹¹ This implies that the conditional short-run momentum strategy outperforms the long-run momentum strategy during high-volatility periods.

Based on the predictions of our model regarding the relative profitability of short-run and long-run momentum strategies, we propose a mixed momentum strategy

11. See Proposition 1 in Daniel, Hirshleifer, and Subrahmanyam (1998).

similar to the one in Goulding, Harvey, and Mazzoleni (2022). Our proposed strategy involves mixing short-run and long-run momentum signals in a way that accounts for the impact of market volatility on the relative profitability of the short-run and long-run momentum strategies.

In particular, we propose the following strategy:

$$\begin{aligned} r_{MM,t} &= w_{t-1}r_{LM,t} + (1 - w_{t-1})r_{SM,t} \\ w_t &= \frac{K}{\sigma_{mkt,t-1}} \end{aligned} \tag{2.21}$$

Here, the weight of the long-run strategy, w_t is inversely related to market volatility and K is constant. We use $r_{LM,t}$ to denote the return of the long-run momentum strategy and $r_{SM,t}$ to denote the return of the short-run momentum strategy.

The $w_t r_{LM,t}$ term represents a volatility-managed momentum strategy, which has been shown to significantly enhance the profitability of conventional 12-2 momentum. In order to limit our focus on the implications of mixing short- and long-run momentum, we also consider a restricted version of the strategy in (2.21):

$$\begin{aligned} r_{MMR,t} &= w'_{t-1}r_{LM,t} + (1 - w'_{t-1})r_{SM,t} \\ w'_{t-1} &= \min\left(\frac{K'}{\sigma_{mkt,t-1}}, 0.999\right) \end{aligned} \tag{2.22}$$

The restricted strategy rules out the possibility of leveraging up and investing heavily in the long-run momentum strategy during low-volatility periods. It also forces us to invest non-zero amounts in the short-run momentum strategy.

Implementation

In order to examine the empirical performance of the mixed momentum strategy, we start by downloading the long-run momentum signal (12-2) from Kenneth French’s data library. Then we construct a short-run momentum signal following the methodology described on Kenneth French’s website. We choose to use the returns of the stocks in months $t - 2$ and $t - 3$ as our short-run momentum signal (3-2). At the end of month t , we sort stocks into deciles based on the short-run momentum signal using NYSE breakpoints. Using the month $t - 2$ return as a momentum signal produces similar results. In Appendix 2.B.2, we reimplement the mixed momentum strategies using the $t = 12$ to $t = 1$ and $t = 3$ to $t = 1$ momentum factors from the Jensen-Kelly-Pedersen data repository.

In order to ensure consistency with the rest of our paper, we use the realized variance of daily market returns during the 126 days preceding the portfolio formation date as a proxy for $\sigma_{mkt,t}$. Using the realized volatility of the market for month $t - 1$ produces similar results.

To implement our strategy, we start with the sample between 01/1950 and 12/2021. We choose 1950 as a start date to avoid the effects of the Great Depression and World War II.

We use the period between 01/1950 and 12/1985 as our training sample and evaluate the performance of the mixed momentum strategy over the period between 01/1986 and 12/2021. The test sample is chosen to match the sample used to test the rationality of equity analysts’ earnings expectations. We use the training sample to pin down the value of the K (K') parameter. We choose the value of K (K') that

maximizes the Sharpe ratio of the unrestricted (restricted) mixed momentum strategy during the training sample. In order to keep our analysis as simple as possible, we use the same K (K') parameter during our entire test sample.

Performance

Figure 2.5 plots the cumulative nominal returns to the unrestricted and restricted mixed momentum strategies compared to the 12-2 momentum strategy over our test sample. We invest \$1 in 1986 and plot the cumulative returns on a log scale for each strategy. The unrestricted and restricted mixed momentum strategies generate about \$57 and \$29, respectively at the end of our sample, compared with about \$10 for the 12-2 momentum strategy.

The restricted momentum strategy follows the returns of the 12-2 strategy very closely for the majority of our sample: the strategy puts more than 90% weight on the long-run momentum signal about 64% of the time. The divergence between the two strategies takes place during the 6% of our sample when the mixed momentum strategy places less than 50% weight on the short-run momentum signal. These high-volatility episodes coincide with the crashes of the 12-2 strategy and the mixed momentum strategy placing less weight on the long-run momentum signal lessens the impact of those crashes, i.e. the returns of the mixed momentum strategy behave in a way that is consistent with the predictions of our model. The unrestricted strategy achieves returns superior to the restricted strategy by taking on relatively more risk when volatility is low, consistent with previous findings on volatility-managed momentum.

Spanning Regressions

To examine the ability of existing trading strategies to account for the profitability of mixed momentum, we estimate time series regressions of mixed momentum on baseline 12-2 momentum, as well as the Fama and French (2015) and Hou, Xue, and Zhang (2014) models, following Moreira and Muir (2017):

$$r_{MM,t} = \alpha + \sum_{j=1}^N \beta_j F_j + \varepsilon_t \quad (2.23)$$

A positive α implies that investors who are already trading the explanatory strategies could realize significant gains by also trading the strategy on the left-hand side of the regression.

The intercepts of these spanning regressions for both the restricted and unrestricted versions of the mixed momentum strategy are reported in Table 2.4. The intercepts are positive and statistically significant in all specifications considered in the table. The two mixed momentum strategies have annualized α 's of about 5.4% (3.7%) relative to the 12-2 momentum strategy. The results reported in columns (2) and (3) show that the state-of-the-art factor models cannot fully account for the returns of the mixed momentum strategy.

2.5.3 Attenuation of momentum

A prediction of the model is that high λ periods are associated with low momentum returns. Additionally, we show that information rigidity displays a secular decline during the period between 1986 and 2021. Therefore, a prediction of our model is

that the profitability of momentum declines over our sample. In the context of our model, the declining cost of information acquisition leads to declining information rigidity and lower profitability of momentum. Figure 2.6 depicts the model-implied relation between the marginal cost of information acquisition (φ) and the profitability of momentum.

In order to test the prediction of the model regarding the declining profitability of momentum, we estimate the following regression, based on Chordia, Subrahmanyam, and Tong (2014):

$$Y_t = ae^{bt+u} \quad (2.24)$$

where Y_t is one plus the return of the 12-2 momentum strategy and t is a time index. We scale the time index to be between -1 and 1 so that the mean of the time variable is zero, as in the original paper. We estimate the regression using momentum return data for the period between 01/1986 and 12/2021 to match the data used to generate Figure 2.1.

The estimate of the b coefficient is -0.014 . It is significant at the 5% level (p-value of one-tailed test 0.025).¹² Since the return of the momentum strategy is positive, a negative coefficient signifies a decline in the profitability of momentum over time, thus providing evidence consistent with the prediction of our model.

12. Following Chordia, Subrahmanyam, and Tong (2014), we test the null hypothesis of no decline in the profitability of momentum.

2.5.4 Return Predictability

2.5.4.1 Return predictability in the model

According to the model, the period 2 expectation of the period 3 price appreciation under the physical measure is:

$$\mathbb{E}_2^P [R_3] = \mathbb{E}_2^P [V - p_2] = \bar{\lambda}_2 \left(\mathbb{E}_2^P [V] - \bar{\mathbb{E}}_2 [V] \right) + \psi x \quad (2.25)$$

where $\bar{\lambda}_2 \equiv \int_0^1 \lambda_{i,2} di$ is the average probability of acquiring the signal in period 2.

The risk premium component of returns, ψx is constant in our model. Therefore, return predictability is driven by the gap between the objective payoff expectations and the average payoff expectations of the market. The magnitude of the gap is determined by the interplay of the two effects, which distort market participants' expectations: rational inattention and overconfidence.

In the context of a multi-period economy, Equation (2.25) states that returns are forecastable using the predictable component of earnings forecast errors. This model prediction carries significant intuitive appeal. If a stock's earnings expectations are overly optimistic, its actual earnings will, on average, fail to meet consensus expectations, which will translate into low returns, i.e. stocks with overly optimistic earnings expectations will deliver low returns. A similar logic applies to stocks with overly pessimistic earnings expectations delivering high returns.

2.5.4.2 Rational earnings expectations

In order to empirically test the model implications regarding return predictability, we need to measure the earnings expectations under the objective measure \mathbb{P} .

The traditional approach in the literature has been to use a linear regression framework to estimate the predictable component of equity analysts' forecast errors (So 2013) or to use the cross-sectional median earnings forecasts as a proxy for objective expectations (La Porta 1996). However, recent work (e.g. Cao and You 2021 and Van Binsbergen, Han, and Lopez-Lira 2022) documents that models that allow for non-linear relations between realized earnings and earnings predictors significantly improve our ability to forecast earnings.

In this paper, we follow Van Binsbergen, Han, and Lopez-Lira (2022) and utilize a tree-based algorithm that can accommodate non-linearities and interactions among the predictors. In particular, we opt for an Extreme Gradient Boosting (XGBoost) algorithm (Chen and Guestrin 2016). XGBoost is chosen due to its speed and superior performance in a large number of settings. Appendix 2.C contains a brief discussion regarding the technical aspects related to XGBoost. In the main body of the paper, we focus our discussion around issues related to the implementation of the algorithm.

Earnings expectations

Let the earnings of firm f during fiscal year $\tau + h$ be

$$e_{f,\tau+h} = \mathbb{E}_t^P [e_{f,\tau+h}] + \varepsilon_{f,t,\tau+h} \quad (2.26)$$

where $\mathbb{E}_t^P [e_{f,\tau+h}] = g(z)$ and z is the $P \times F \times T$ -dimensional matrix of predictors. We assume that $g(\cdot)$ is a flexible function of these predictors. We only impose the restriction that $g(\cdot)$ does not depend on f or t , i.e. the function is the same for all firms and over time.

Tree-based methods

Our goal in this section is to use XGBoost to approximate the function $g(\cdot)$.

XGBoost is based on decision trees in which the data is recursively split into non-intersecting partitions. The algorithm approximates $g(\cdot)$ with the average value of the outcome variable in each partition. At each step, the algorithm groups observations that behave similarly by minimizing the mean squared error when the average value of the dependent variable in each partition is used to form forecasts. Due to the large number of potential splits, tree-based methods rely on "greedy" optimization, which involves myopically minimizing forecast errors during each split.

The $g(\cdot)$ for a tree with K terminal nodes (leaves) can be formally written as:

$$\mathbb{E}^P [y] = g(z) = \sum_{\kappa=1}^K x_{\kappa} \mathbf{1}_{\{z \in C_{\kappa}\}} \quad (2.27)$$

where x_{κ} is the sample average of the dependent variable in partition κ and is given by:

$$x_{\kappa} = \frac{1}{N_{\kappa}} \sum_{y: z_p \in C_{\kappa}} y \quad (2.28)$$

and region C_κ is chosen by forming hyper-regions in the space of predictors:

$$C_\kappa = \left\{ z_p \in \times_{p \in P} Z_p : \underline{z}_p^\kappa < z_p \leq \bar{z}_p^\kappa \right\} \quad (2.29)$$

where \times denotes Cartesian product, P is the number of predictors, and each predictor z_p can take values in set Z_p .

The decision tree in Figure 2.7 illustrates the contents of Equations (2.27), (2.28), and (2.29) using a simple example. The outcome variable in this example is EPS and the predictors we consider are lagged EPS realizations, lagged prices, and lagged returns. Given the structure in the figure, $g(\cdot)$ takes on the following form:

$$\begin{aligned} g(z; y) = & (0.40)\mathbf{1}_{\{past\ EPS \leq 1\}} + (1.70)\mathbf{1}_{\{past\ EPS > 1\}}\mathbf{1}_{\{past\ price \leq 20\}} + \\ & + (2.14)\mathbf{1}_{\{past\ EPS > 1\}}\mathbf{1}_{\{past\ price > 20\}}\mathbf{1}_{\{past\ return \leq 10\%\}} + \\ & + (3.50)\mathbf{1}_{\{past\ EPS > 1\}}\mathbf{1}_{\{past\ price > 20\}}\mathbf{1}_{\{past\ return > 10\%\}} \end{aligned} \quad (2.30)$$

Extreme Gradient Boosting

XGBoost is an algorithm that is based on recursively combining forecasts from a large number of weak learners to form a strong learner. During the first step of implementation, the algorithm fits a weak learner to the training sample. At each subsequent step s , the algorithm fits a weak learner to the residuals of a model with $s - 1$ trees. The residual forecast is then added to the total with a shrinkage weight $\eta \in (0, 1)$. The additional forecasts are shrunk to avoid overfitting the residuals.

In order to implement the XGBoost algorithm, four hyper-parameters need to be chosen: γ , **maximum depth**, and **subsample** in addition to the already-described shrinkage parameter η . Hyper-parameters are a characteristic of a model whose value cannot be estimated from data. Therefore, we tune (choose the values of) the hyper-parameters using a cross-validation procedure that we describe later in the paper. The parameter values we use throughout our analysis are presented in Table 2.3.

The γ parameter determines the minimum loss reduction required to make a further partition on a leaf node of the tree. This parameter controls the total number of trees in the ensemble.

Maximum depth determines the maximum depth (complexity) of the tree. More complex models are more likely to overfit the training sample.

Subsample determines the ratio of training instance. A value of 0.15 means that XGBoost randomly collects 15% of the observations to use as a training sample for each decision tree.

Earnings forecasts

In our analysis, we focus on two-year earnings forecasts ($FPI = 2$). The long forecasting horizon maximizes the scope for expectation errors and increases the chances of uncovering interesting asset pricing dynamics. Van Binsbergen, Han, and Lopez-Lira (2022) find that the bias of equity analysts' expectations increases with the forecasting horizon and Silva and Thesmar (2023) find that analysts' forecasts outperform statistical forecasts for forecasting horizons of less than one year. To match the frequency of I/B/E/S analyst forecasts, we construct objective expecta-

tions for each month.

We deviate from existing papers that forecast earnings by only using variables that are available through CRSP (prices and returns) or I/B/E/S (earnings forecasts and past EPS realizations). We choose this limited set of predictors to avoid basing expectations on information not available to the equity analysts in real time. Variables extracted from financial statements, which have been shown to predict future earnings, may be restated after initial publication and accounting restatements affect stock prices (Hribar and Jenking 2004 and Palmrose, Richardson, and Scholz 2004).¹³ Therefore, using variables extracted from financial statements may contaminate the out-of-sample tests presented in this paper.

An additional advantage of restricting our attention to a limited number of earnings predictors is that by doing so we sidestep the missing data problem outlined in Bryzgalova et al. (2022). Using the readily-available predictors allows to avoid having to take a stand regarding the appropriate approach to filling in missing firm characteristics.

Using the notation established in this section, the objective earnings expectations are represented as:

$$\mathbb{E}_t^P [e_{f,\tau+1}] = g(e_{f,\tau-1}, \mathbb{F}_t[e_{f,\tau+1}], P_{f,t}, r_{f,t}) \quad (2.31)$$

To implement the XGBoost algorithm, we split our sample into three non-overlapping time periods: training subsample, validation subsample, and testing subsample. The training sample covers the period between 1976 and 1983 and is used to estimate the

13. Realized earnings in I/B/E/S are not restated.

model using a set of hyper-parameter values.

Our validation sample encompasses the period between 01/1984 and 11/1985 and is used to conduct quasi-out-of-sample tests: we construct predicted earnings for 1986 and 1987 based on our model and compute the mean squared error corresponding to the set of hyper-parameters used to train the model. Next, we reestimate the model using a different set of hyper-parameters. We grid-search over various hyper-parameter combinations and select the combination that minimize the out-of-sample mean squared error of our model. Once the hyper-parameters are chosen, we use the same parameters for all of our out-of-sample tests.

We start our out-of-sample forecasts in 11/1985 and use ten-year rolling windows to train the model. The algorithm provides us with two-year earnings forecasts for the period between 12/1985 and 11/2019. The objective forecasts do not rely on information that is not available to the market participants by the end of month t . Therefore, the trading strategy proposed in the next subsection is implementable in real time.

2.5.4.3 Return predictability in the data

The XGBoost algorithm outlined in the previous subsection provides us with objective two-year earnings forecasts for each month t between 12/1985 and 11/2019. Given the objective forecasts, we can compute the predictable component of analysts' expectation errors using the following formula:

$$\widehat{\text{EE}}_{f,t}^{\tau+1} = \frac{\mathbb{E}_t^P [e_{f,\tau+1}] - \mathbb{F}_t [e_{f,\tau+1}]}{P_{f,t-1}} \quad (2.32)$$

Negative values of \widehat{EE} indicate that analysts' expectations are excessively optimistic and positive values of the measure indicate excessive pessimism. Following Engelberg, McLean, and Pontiff (2018), the difference between the expectations under the objective and subjective forecasts is normalized by lagged stock prices to ensure greater comparability across firms.

In order to test the predictions of the model regarding return predictability, we sort stocks into quintiles based on the value of \widehat{EE} at the end of month t . Panel A of Table 2.5 reports the one-month holding period returns of the five portfolios. The results reported in the table are for the period between 01/1986 and 12/2019.

The results in Panel A of Table 2.5 support the predictions of the model: the value-weighted portfolio returns increase monotonically in \widehat{EE} . In particular, a portfolio that is long on stocks in the fifth quintile and short on stocks in the first quintile earns an average return of 1.59% per month. For the rest of this paper, we refer to this long-short trading strategy as pessimistic-minus-optimistic or PMO. The t -statistic testing whether the PMO premium is zero is 6.82. Thus, PMO clears the hurdle of a t -statistic ≥ 3.0 proposed by Harvey, Liu, and Zhu (2015). Monthly returns of 1.59% seem large in comparison to existing trading strategies (for comparison, the average value-weighted market return between 01/1986 and 12/2019 is about 1%). However, the profitability of our strategy is in line with the profitability of the strategies considered by Van Binsbergen, Han, and Lopez-Lira (2022).

In order to assess the degree to which PMO's profitability is driven by the smallest and most illiquid stocks within our sample, we construct an alternative long-short trading strategy by restricting our investment opportunity set to only include stocks whose market capitalization is above the 90th percentile of market capitalizations in a

given year.¹⁴ The results of the trading strategy that invests only in large-cap stocks are reported in Panel B of Table 2.5. Unsurprisingly, the results in this instance are somewhat weaker than those reported in Panel A but the return of the long-short portfolio is still about 0.77% (t -statistic = 4.74) per month. Implementing the PMO trading strategy using only the most liquid stocks in our sample generates significant positive returns.

In Panel C, we examine if the profitability of the PMO strategy is driven by the earlier years within our sample. In this test, we exclude the years between 1986 and 2002 from our sample. The return of the long-short portfolio for the period between 01/2003 and 12/2019 is about 1.53% per month (t -statistic = 5.23) and we are unable to reject the hypothesis that the returns during the second half of our sample are equal to the returns during the first half of our sample (t -statistic = -0.34).

Spanning regressions

We further use spanning regressions to assess the ability of traditional factor models to explain the profitability of the PMO strategy. Specifically, we estimate the following regressions and examine the significance of the regression intercepts (alphas):

$$PMO_t = \alpha + \sum_{j=1}^N \beta_j F_{j,t} + \varepsilon_t \quad (2.33)$$

where $F_{j,t}$, $j = \{1, 2, 3, \dots, N\}$ are the excess market return, the five Fama and French (2015) factors augmented with the WML factor, or the four factors of the q-factor

14. The minimum market capitalization of a firm included in the megacap sample is about 2 billion in the early parts of the sample and over 21 billion near the end of the sample.

model (Hou, Xue, and Zhang 2014). The results of the spanning regressions are reported in Table 2.6.

The full-sample tests in Panel A of Table 2.6 show that PMO earns a significant alpha with t -statistics greater than 5.0 relative to the three models considered in the table. This suggests that the factors considered in the table are unable to fully account for the profitability of the PMO strategy. In Panels B and C we verify that our conclusions regarding the inability of factor models to account for the profitability of PMO hold in both the first half (1986-2002) and the second half (2003-2019) of our sample. It is noteworthy that the performance of PMO improves significantly relative to the explanatory trading strategies in the second half of our sample.

In unreported tests, we show that the profitability of the PMO strategy cannot be explained using a combination of popular trading strategies proposed in the literature. While an explanation of the profitability of the PMO strategy is beyond the scope of this paper, we believe that the PMO strategy should receive a rigorous treatment in the literature going forward.

2.6 Conclusion

In this paper, we document the existence of time variation in the stickiness of financial analysts' expectations. The stickiness of analysts' expectations declines during high-volatility periods. Additionally, expectation stickiness experiences a sustained decline over our sample.

To account for these stylized facts, we build a simple featuring time-varying inattention. We explore the asset pricing implications of our model and show that

it is consistent with positive unconditional momentum returns, momentum crashes, the profitability of volatility-managed momentum, and the diminishing profitability of momentum over our sample.

In order to test our model's prediction regarding return predictability, we extract the predictable component of analysts' forecast errors and propose a trading strategy that is long (short) on stocks with excessively pessimistic (optimistic) earnings forecasts. Existing prominent factor models cannot fully explain the profitability of our trading strategy, especially during the second half of our sample.

Table 2.1
Forecast Error Predictability Regressions

	<i>Dependent variable:</i>						
	$FE_{f,\tau}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	RV				IV		
$FR_{f,t}$	0.167*** [5.195]	0.243*** [5.497]	0.201*** [7.230]	0.197*** [7.546]	0.285*** [4.817]	0.195*** [7.280]	0.199*** [7.722]
$FR_{f,t} \cdot \hat{\sigma}_{\text{mkt}}$		-0.065 [-1.541]			-0.056* [-1.848]		
$FR_{f,t} \cdot \mathbf{1}_{\text{HV}}$			-0.126* [-1.900]	-0.160* [-1.848]		-0.128* [-1.984]	-0.190** [-2.288]
Constant	-0.012*** [-8.190]	-0.013*** [-5.942]	-0.011*** [-8.438]	-0.012*** [-8.782]	-0.007** [-2.603]	-0.011*** [-7.739]	-0.011*** [-8.436]
$\hat{\sigma}_{\text{mkt}}$		0.001 [0.502]			-0.002 [-1.404]		
$\mathbf{1}_{\text{HV}}$			-0.001 [-0.236]	0.001 [0.340]		-0.005 [-1.634]	-0.002 [-0.420]
Observations	78,287	78,287	78,287	78,287	77,969	77,969	77,969
$P(\mathbf{1}_{\text{HV}} = 1)$	—	—	0.20	0.10	—	0.20	0.10
Adjusted R ²	0.014	0.015	0.015	0.016	0.015	0.015	0.016

This table reports the results for the forecast error predictability regression $FE_{f,\tau} = \alpha^{CG} + \beta^{CG} FR_{f,t} + \varepsilon_{t,\tau}$ and the modified regression $FE_{f,\tau} = \alpha^{CG} + \alpha_{\Delta}^{CG} z_t + (\beta^{CG} + \beta_{\Delta}^{CG} z_t) FR_{f,t} + \varepsilon_{t,\tau}$, where z_t in columns (2) and (5) is realized and implied volatility, respectively, and a high volatility indicator in columns (3), (4), (6), and (7). Realized volatility is based on daily value-weighted CRSP returns for a period of 126 days. The CBOE S&P100 Volatility Index is used as a measure of implied volatility. $FE_{f,\tau}$ is defined as $(e_{f,\tau} - \mathbb{F}_t[e_{f,\tau}]) / P_{f,t-1}$ and $FR_{f,t}$ is defined as $(\mathbb{F}_t[e_{f,\tau}] - \mathbb{F}_{t-1}[e_{f,\tau}]) / P_{f,t-1}$. Standard errors are double-clustered by firm and year. The corresponding t -statistics are in square brackets. Significance at the 1%, 5%, and 10% is denoted by ***, **, and *, respectively. Earnings forecasts are obtained from I/B/E/S and cover firms with fiscal year ends between 1/1986 and 11/2021. The volatility measures are in percent per month.

Table 2.2
Momentum and Volatility

	<i>Dependent variable:</i>						
	$r_{\text{WML},t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	RV				IV		
$\hat{\sigma}_{\text{mkt}}$		-2.839** [-2.038]			-1.750** [-1.977]		
$\mathbf{1}_{\text{HV}}$			-0.026* [-1.805]	-0.036 [-1.459]		-0.028** [-2.077]	-0.066*** [-3.092]
Constant	0.009** [2.207]	0.037*** [2.930]	0.014*** [3.983]	0.012*** [3.459]	0.038*** [2.911]	0.014*** [4.586]	0.015*** [4.564]
Observations	432	432	432	432	422	422	422
$P(\mathbf{1}_{\text{HV}} = 1)$	—	—	0.20	0.10	—	0.20	0.10

This table reports the results for the following regression: $r_{\text{WML},t} = \alpha + \alpha_{\Delta} z_{t-1} + \varepsilon_t$. In columns (2) and (5) z_{t-1} is realized and implied volatility, respectively. In columns (3), (4), (6), and (7) z_{t-1} is a high volatility indicator. Realized volatility is based on daily value-weighted CRSP returns for a period of 126 days. The CBOE *S&P*100 Volatility Index is used as a measure of implied volatility and the volatility of value-weighted CRSP is used as a measure of realized volatility. The standard errors are computed using the Newey and West (1987) methodology with six lags. The corresponding t -statistics are in square brackets. Significance at the 1%, 5%, and 10% is denoted by ***, **, and *, respectively. The momentum return data is obtained from Kenneth French's website and covers the period between 01/1986 and 12/2021.

Table 2.3
XGBoost Hyper-parameters

η	0.01
γ	0.15
maximum depth	7
subsample	0.15
nrounds	10000

This table reports the hyper-parameters chosen for the XGBoost algorithm.

Table 2.4
Mixed Momentum Spanning Regressions

	Momentum	FF6	q-Factor
Panel A: Unrestricted			
$\hat{\alpha}$	0.45	0.68	0.89
t -statistic	3.04	3.16	2.25
Panel B: Restricted			
$\hat{\alpha}$	0.31	0.48	0.70
t -statistic	2.87	2.82	2.02

This table reports the $\hat{\alpha}$ (in %) obtained by estimating the following regression: $r_{MM,t} = \alpha + \sum_{j=1}^N \beta_j F_{j,t} + \varepsilon_t$ where the factors are: the return on the 12-2 momentum strategy, the five Fama-French factors + WML, or the four Hou-Xue-Zhang factors. The full sample covers the period between 01/1986 and 12/2021. The t -statistics are based on standard errors computed following White (1980).

Table 2.5
Portfolios Sorted on Expectation Errors

Quintile	1	2	3	4	5	5 – 1
Panel A: All Stocks						
Mean	–0.11	0.59	0.88	1.03	1.48	1.59
<i>t</i> -statistic	–0.32	2.19	3.77	5.10	7.08	6.82
Panel B: Market cap above 90 th percentile						
Mean	0.56	0.90	0.85	1.07	1.33	0.77
<i>t</i> -statistic	2.04	3.95	4.00	5.47	6.15	4.74
Panel C: Post-2002 Sample						
Mean	0.15	0.54	0.77	0.94	1.68	1.53
<i>t</i> -statistic	0.30	1.59	2.59	3.83	6.31	5.23

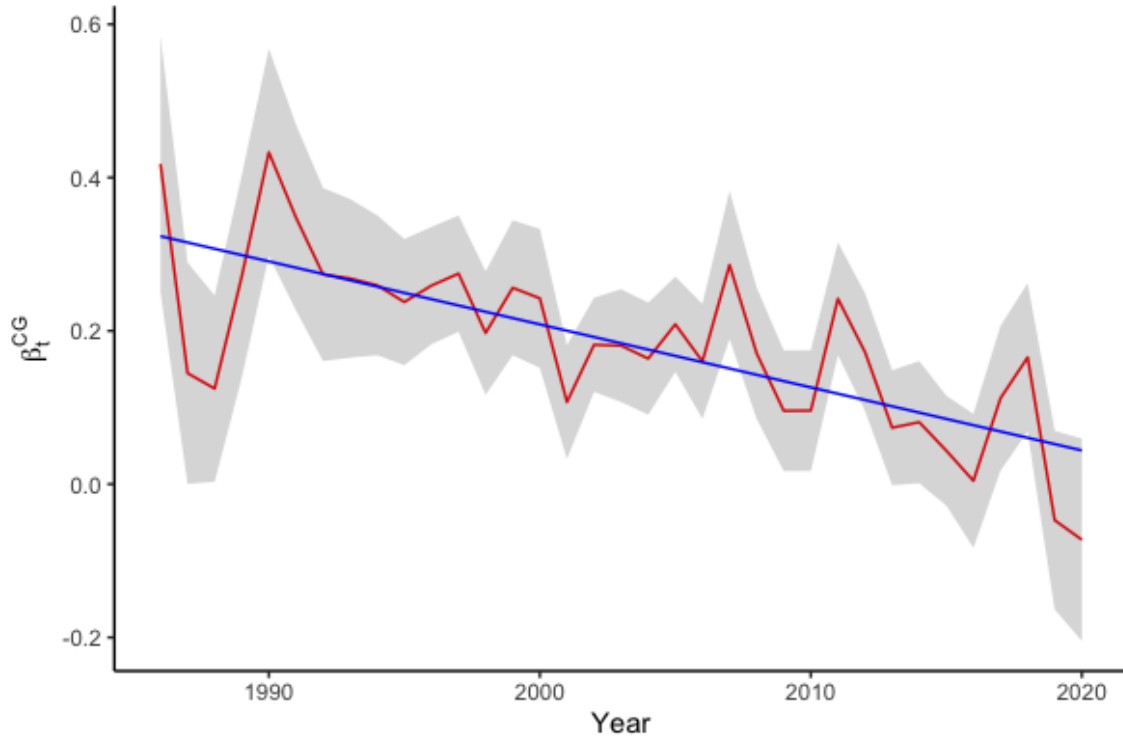
This table reports the time-series average returns on value-weighted portfolios formed based on the predictable component of equity analysts' forecast errors, \widehat{EE} . The full sample in Panel A includes all stocks and covers the period between 01/1986 and 12/2019. In Panel B, the sample is restricted to stocks with market caps above the 90th percentile of market caps within a given year. The sample in Panel C covers the period between 01/2003 and 01/2020.

Table 2.6
PMO Spanning Regressions

	CAPM	FF6	q-Factor
Panel A: Full Sample			
$\hat{\alpha}$	1.90	1.51	1.63
t -statistic	8.97	8.78	6.30
Panel B: 01/1986-12/2002			
$\hat{\alpha}$	1.81	1.22	1.41
t -statistic	5.27	5.55	2.92
Panel C: 01/2003-01/2020			
$\hat{\alpha}$	2.08	1.68	1.87
t -statistic	8.91	8.88	8.00

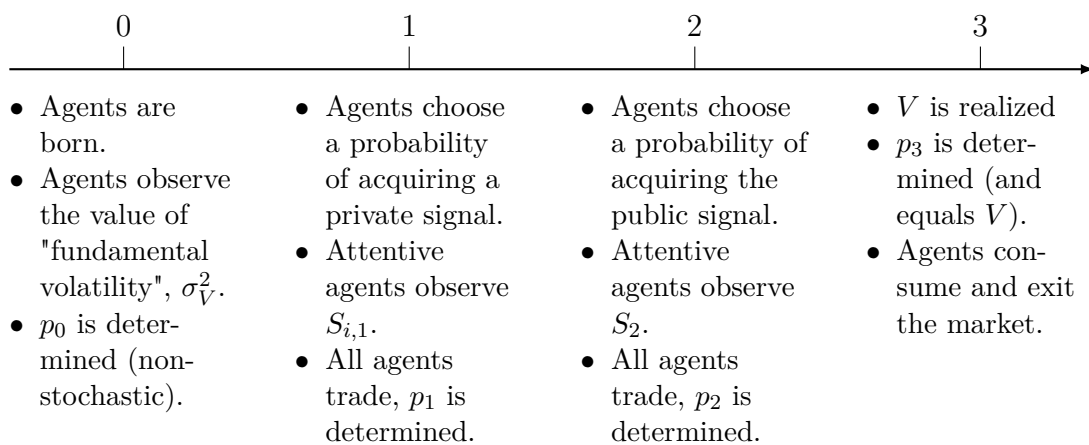
This table reports the $\hat{\alpha}$ (in %) obtained by estimating the following regression: $PMO_t = \alpha + \sum_{j=1}^N \beta_j F_{j,t} + \varepsilon_t$ where the factors are: excess market return, the five Fama-French factors + WML, or the four Hou-Xue-Zhang factors. The full sample covers the period between 01/1986 and 12/2019. The t -statistics are based on standard errors computed following White (1980).

Figure 2.1
Information Rigidity Over Time



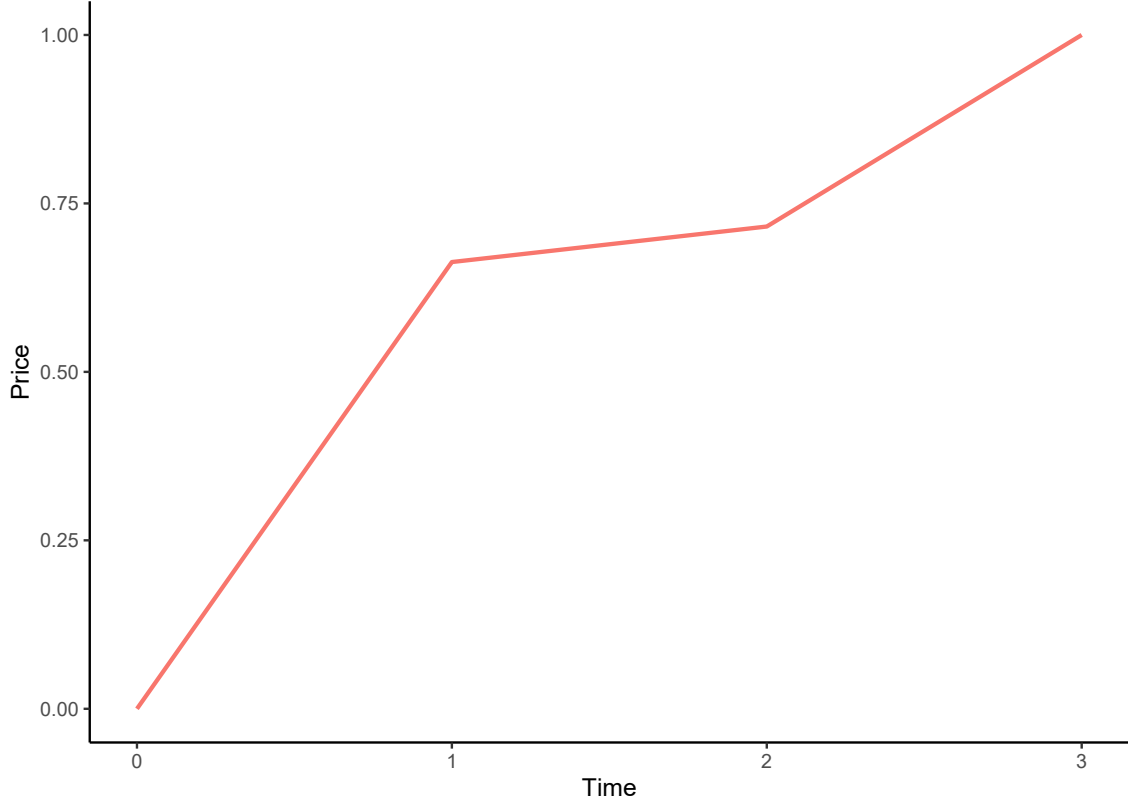
The figure above depicts the coefficient β_t^{CG} on forecast revisions in specification (2.1) estimated using overlapping two-year windows. The shaded region represents the 95% confidence interval. The estimation is carried out using data from I/B/E/S and covers firms with fiscal year ends between 01/1986 and 12/2021.

Figure 2.2
Model Representation



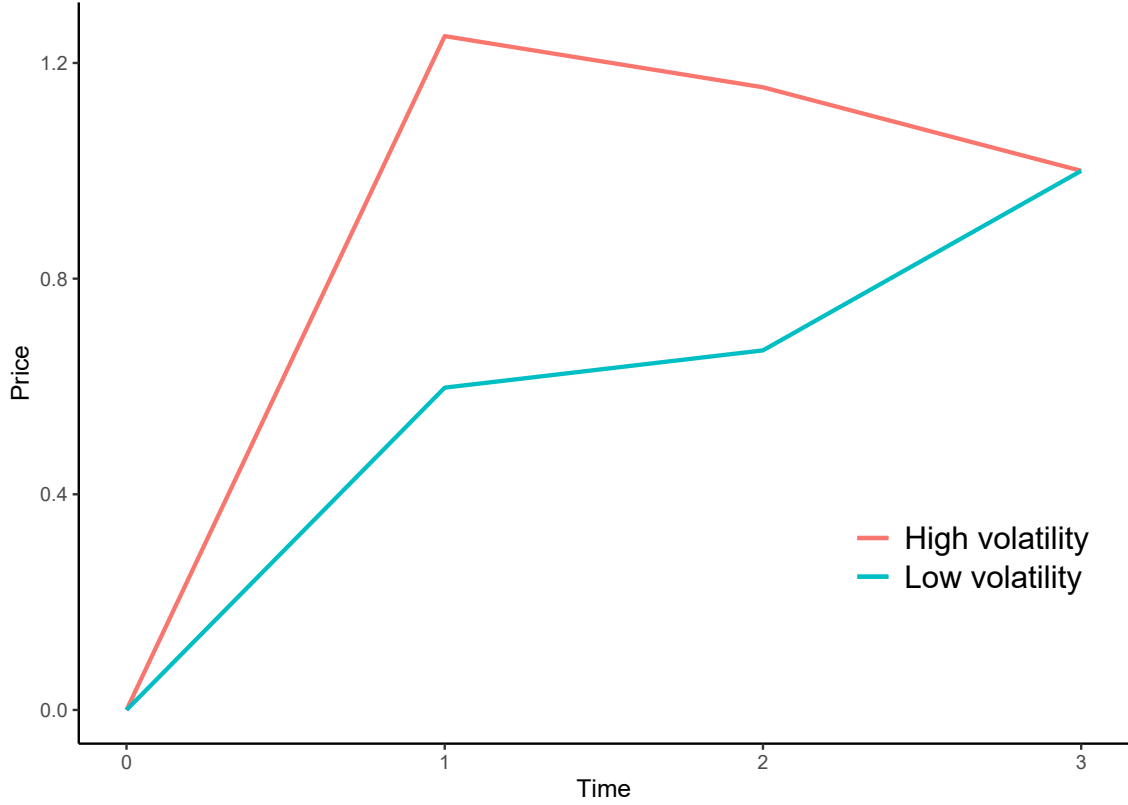
This figure summarizes the life cycle of an agent born in period 0.

Figure 2.3
Average Price Path



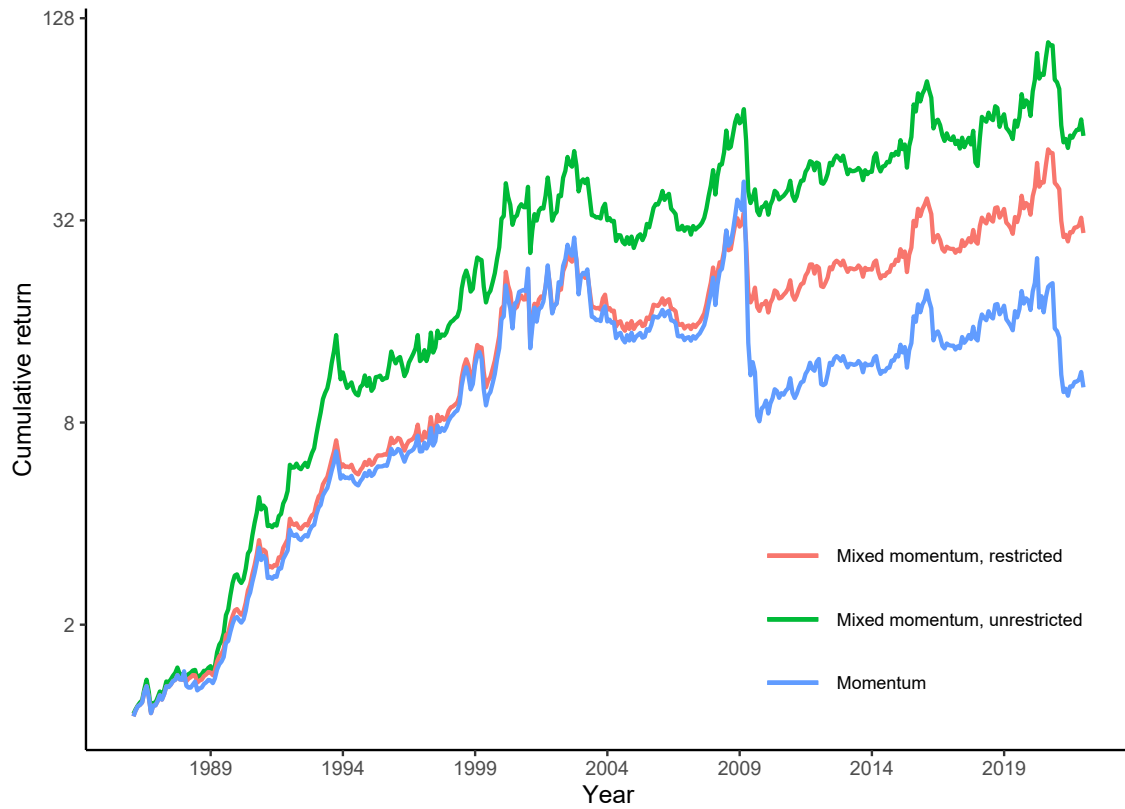
The figure above depicts the price paths for the model in Section 2.4. The x-axis represents time and the y-axis is the price scaled by fundamental value V . The figure is based on the following parameters: $\pi_{hv} = 0.10$, $\sigma_{V,lv}^2 = 0.35$, $\sigma_{V,hv}^2 = 0.60$, $\sigma_S^2 = 0.70$, $\sigma_C^2 = 0.30$, $\psi = 1$, $x = 10^{-5}$, $\varphi = 0.70$, and $\kappa = 4$.

Figure 2.4
Average Price Paths, High and Low Volatility States



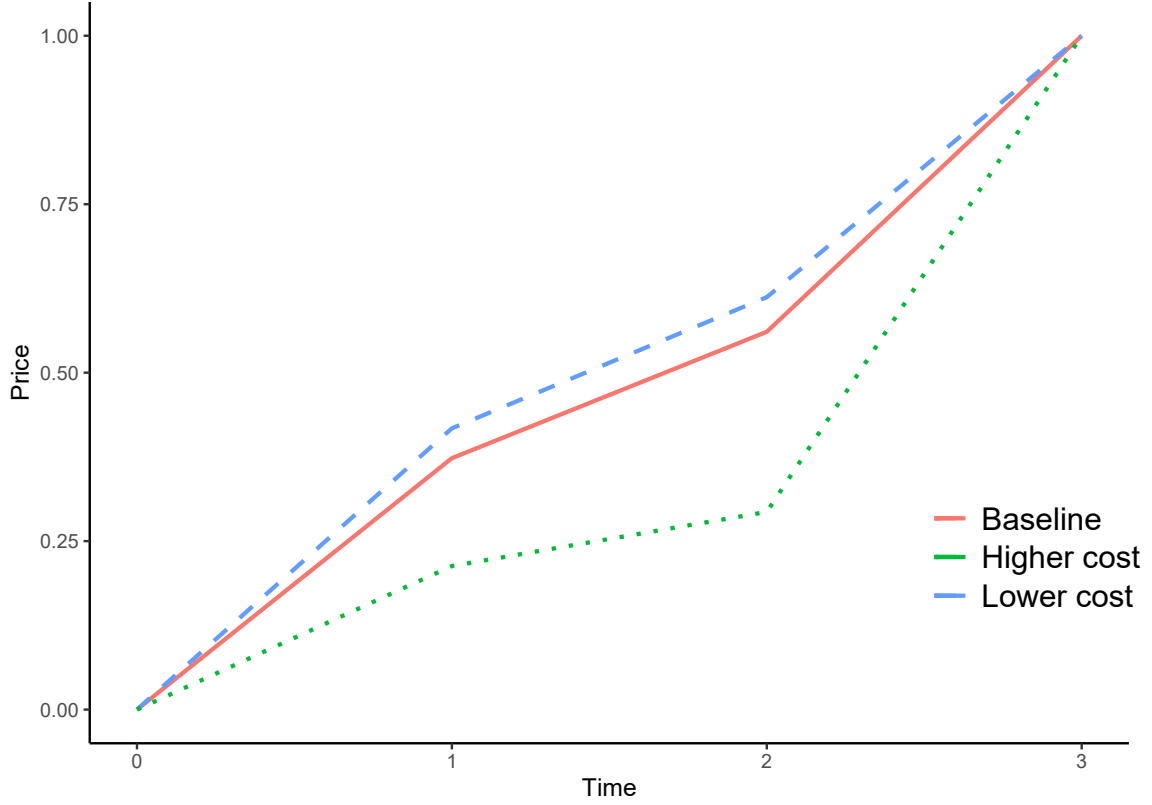
The figure above depicts the price paths based on the model in Section 2.4 for high volatility and low volatility states. The x-axis represents time and the y-axis is the price scaled by fundamental value V . The figure is based on the following parameters: $\sigma_{V,lv}^2 = 0.35$, $\sigma_{V,hv}^2 = 0.60$, $\sigma_S^2 = 0.70$, $\sigma_C^2 = 0.30$, $\psi = 1$, $x = 10^{-5}$, $\varphi = 0.70$, and $\kappa = 4$.

Figure 2.5
Mixed Momentum Strategy



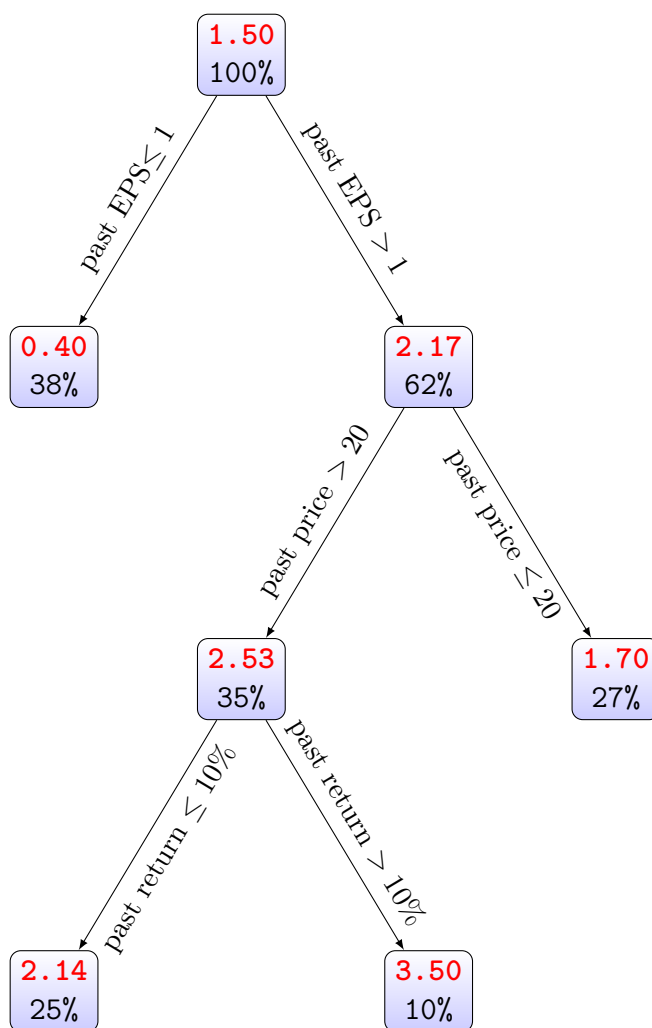
The figure above depicts the cumulative returns of the restricted and unrestricted versions of the mixed momentum trading strategy developed in Section 2.5.2.1, along with the returns of the baseline momentum strategy. In this figure, we consider the time period between 01/1986 and 12/2021.

Figure 2.6
Cost of Information and Momentum



The figure above depicts the price paths based on the model in Section 2.4 for different marginal costs of information acquisition. The x-axis represents time and the y-axis is the price scaled by fundamental value V . The figure is based on the following parameters: $\sigma_V^2 = 0.35$, $\sigma_S^2 = 0.70$, $\sigma_C^2 = 0.30$, $\psi = 1$, $x = 10^{-5}$, $\varphi_{\text{baseline}} = 0.70$, $\varphi_{\text{low cost}} = 0.50$, $\varphi_{\text{high cost}} = 0.90$ and $\kappa = 4$.

Figure 2.7
Decision Tree Example



This figure graphically illustrates the structure of a decision tree. Our goal is to predict EPS and we use past EPS, past price, and past return as predictors. The percentages represent the proportion of our observations that end up in each node and the numbers in red represent the average EPS of the stocks within each node.

APPENDICES

2.A Derivations

2.A.1 Information Acquisition Problem

Our goal is to obtain an expression for $\mathbb{E}_{t-1}^i [U_{i,t}]$.

First, we plug in the expression for optimal $q_{i,t}$ and obtain the following maximization problem:

$$\mathbb{E}_{t-1}^i \left[\frac{\mathbb{E}_t^i[V] - p_t}{\psi} (\mathbb{E}_t^i[V] - p_t) - \frac{\psi (\mathbb{E}_t^i[V] - p_t)^2}{2\psi^2} \right] = \frac{1}{2\psi} \mathbb{E}_{t-1}^i \left[(\mathbb{E}_t^i[V] - p_t)^2 \right] \quad (2.34)$$

We expand the square:

$$\frac{1}{2\psi} \mathbb{E}_{t-1}^i \left[(\mathbb{E}_t^i[V])^2 - 2p_t \mathbb{E}_t^i[V] + p_t^2 \right] \quad (2.35)$$

The price p_t is only affected by aggregate information choices, therefore each agent i takes the price as given (as in Kacperczyk, Van Nieuwerburgh, and Veldkamp 2016). The agents taking prices as given, combined with the fact that the law of iterated expectations holds at the individual level, allows us to rewrite the expression above as:

$$\frac{1}{2\psi} \left(\mathbb{E}_{t-1}^i \left[(\mathbb{E}_t^i[V])^2 \right] - 2p_t \mathbb{E}_{t-1}^i[V] + p_t^2 \right) \quad (2.36)$$

In this expression, $\mathbb{E}_{t-1}^i [(\mathbb{E}_t^i[V])^2]$ is the only term that depends on $\lambda_{i,t}$. Therefore,

we need to evaluate the following expression:

$$\frac{1}{2\psi} \mathbb{E}_{t-1}^i \left[\left(\mathbb{E}_t^i[V] \right)^2 \right] \quad (2.37)$$

We use the definition of variance to rewrite the expression as:

$$\frac{1}{2\psi} \left(\text{var}_{t-1}^i \left(\mathbb{E}_t^i[V] \right) + \left(\mathbb{E}_{t-1}^i \left[\mathbb{E}_t^i[V] \right] \right)^2 \right) \quad (2.38)$$

The second term does not depend on $\lambda_{i,t}$, so we focus on the variance term:

$$\frac{1}{2\psi} \left(\text{var}_{t-1}^i \left(\lambda_{i,t} \mathbb{E}_t[V] + (1 - \lambda_{i,t}) \mathbb{E}_{t-1}[V] \right) \right) = \frac{\lambda_{i,t}^2}{2\psi} \text{var}_{t-1}^i \left(\mathbb{E}_t[V] \right) \quad (2.39)$$

2.A.2 Information Rigidity Coefficient

The model-implied version of the forecast error predictability coefficients is:

$$\beta^{CG} = \frac{\text{cov}(V - p_2, p_2 - p_1)}{\text{var}(p_2 - p_1)}. \quad (2.40)$$

The numerator and denominator in equation (2.40) are given by

$$\begin{aligned} \text{cov}(V - p_2, p_2 - p_1) = & \frac{\bar{K}}{D} \left[\lambda_1^2 \lambda_{2|A} (\sigma_V^2 + \sigma_S^2) \left(D \sigma_V^2 \sigma_C^2 (\sigma_S^2 - \sigma_C^2) \right. \right. \\ & - \lambda_{2|A} \sigma_V^2 \sigma_C^4 (\sigma_C^4 + 2\sigma_V^2 \sigma_S^2 + \sigma_S^2 \sigma_C^2) \Big) \\ & + (1 - \lambda_1) \lambda_{2|I} D^2 (\sigma_V^2 + \sigma_C^2) \left((\sigma_V^2 + \sigma_C^2) \right. \\ & - 2\lambda_1 \lambda_{2|A} \sigma_C^2 - \lambda_1 \sigma_V^2 - (1 - \lambda_1) \lambda_{2|I} (\sigma_V^2 + \sigma_C^2) \Big) \\ & \left. \left. + \lambda_1 \lambda_{2|A} D (\sigma_V^2 + \sigma_S^2) (\sigma_V^2 + \sigma_C^2) \sigma_C^4 \right] \end{aligned} \quad (2.41)$$

$$\begin{aligned} \text{var}(p_2 - p_1) = & \frac{\bar{K}}{D} \left[\lambda_1^2 \lambda_{2|A}^2 (2\sigma_V^4 \sigma_S^2 + 2\sigma_V^2 \sigma_S^2 \sigma_C^2 + \sigma_V^2 \sigma_C^4 + \sigma_S^2 \sigma_C^4) (\sigma_V^2 + \sigma_S^2) \sigma_C^4 \right. \\ & + 2\lambda_1 (1 - \lambda_1) \lambda_{2|A} \lambda_{2|I} D^2 (\sigma_V^2 + \sigma_C^2) \sigma_C^2 \\ & \left. + (1 - \lambda_1)^2 \lambda_{2|I}^2 D^2 (\sigma_V^2 + \sigma_C^2)^2 \right] \end{aligned} \quad (2.42)$$

where $D = \sigma_V^2 \sigma_S^2 + \sigma_V^2 \sigma_C^2 + \sigma_S^2 \sigma_C^2$, and $\bar{K} = \frac{\sigma_V^4}{D(\sigma_V^2 + \sigma_S^2)(\sigma_V^2 + \sigma_C^2)^2}$.

2.A.3 Comparative Statics

The comparative statics of the probability of information acquisition with respect to σ_V^2 are shown below:

$$\frac{\partial \lambda_1}{\partial \sigma_V^2} = \frac{\sigma_V^2 (2\sigma_C^2 \sigma_S^2 + 3\sigma_C^2 \sigma_V^2 + \sigma_V^4)}{\psi \varphi (\sigma_V^2 + \sigma_C^2)^3} > 0 \quad (2.43)$$

$$\frac{\partial \lambda_{2|I}}{\partial \sigma_V^2} = \frac{\sigma_V^2 (2\sigma_S^2 + \sigma_V^2)}{\psi \varphi (\sigma_V^2 + \sigma_S^2)^2} > 0 \quad (2.44)$$

$$\frac{\partial \lambda_{2|I}}{\partial \sigma_V^2} = \frac{1}{\psi \varphi(\sigma_C^2(\sigma_S^2 + \sigma_V^2) + \sigma_S^2 \sigma_V^2)^3} \left[\sigma_V^2(\sigma_C^6(\sigma_S^2 + \sigma_V^2)(2\sigma_S^2 + \sigma_V^2) + \sigma_C^4 \sigma_S^2 \sigma_V^4 \right. \\ \left. + \sigma_C^2 \sigma_V^4(\sigma_S^2 + \sigma_V^2)(4\sigma_S^2 + 3\sigma_V^2) + \sigma_S^2 \sigma_V^6(2\sigma_S^2 + 3\sigma_V^2)) \right] > 0 \quad (2.45)$$

2.A.4 Momentum

The short-run momentum parameter is defined as

$$\text{MOM}_S = \frac{\text{cov}(p_2 - p_1, p_1 - p_0) + \text{cov}(V - p_2, p_2 - p_1)}{2} \quad (2.46)$$

The covariances from equation (2.46) are:

$$\text{cov}(p_2 - p_1, p_1 - p_0) = \lambda_1 \sigma_V^2 \bar{K} \left[(1 - \lambda_1) \lambda_{2|I} D (\sigma_V^2 + \sigma_C^2) \right. \\ \left. - \lambda_1 \lambda_{2|A} \sigma_C^2 (\sigma_V^2 + \sigma_S^2) (\sigma_S^2 - \sigma_C^2) \right] \quad (2.47)$$

where D and \bar{K} are defined in section 2.A.2, and $\text{cov}(V - p_2, p_2 - p_1)$ is in equation (2.42). Therefore, the unconditional short-run momentum parameter MOM_S from equation (2.46) is

$$\text{MOM}_S = \frac{\bar{K}}{2D} \left[-\lambda_1^2 \lambda_{2|A}^2 (\sigma_C^4 + 2\sigma_V^2 \sigma_S^2 + 2\sigma_S^2 \sigma_C^2) (\sigma_V^2 + \sigma_S^2) \sigma_V^2 \sigma_C^4 \right. \\ \left. + (1 - \lambda_1) \lambda_{2|I} D^2 (\sigma_V^2 + \sigma_C^2) \left((\sigma_V^2 + \sigma_C^2) - 2\lambda_1 \lambda_{2|A} \sigma_C^2 \right. \right. \\ \left. \left. - (1 - \lambda_1) \lambda_{2|I} (\sigma_V^2 + \sigma_C^2) \right) + \lambda_1 \lambda_{2|A} D (\sigma_V^2 + \sigma_S^2) (\sigma_V^2 + \sigma_C^2) \sigma_C^4 \right] \quad (2.48)$$

The long-run momentum parameter is

$$\begin{aligned} \text{cov}(V - p_2, p_2 - p_0) = & \bar{K} \left[(1 - \lambda_1) D(\sigma_V^2 + \sigma_C^2) \left((1 - \lambda_{2|I}) \sigma_V^2 + \sigma_S^2 \right) \right. \\ & \left. - \lambda_1 \left(\sigma_V^2 \sigma_S^2 + (1 - \lambda_{2|A}) \sigma_V^2 \sigma_C^2 + \sigma_S^2 \sigma_C^2 \right) (\sigma_V^2 + \sigma_S^2) (\sigma_S^2 - \sigma_C^2) \right]. \end{aligned} \quad (2.49)$$

2.B Robustness: Jensen-Kelly-Pedersen Momentum Factor

2.B.1 Momentum and Volatility

In this section, we estimate the same regression we estimated in Section 2.5.2:

$$r_{WML,t} = \alpha + \alpha_\Delta z_{t-1} + \varepsilon_t \quad (2.50)$$

where $r_{WML,t}$ are the time t momentum returns and z_{t-1} is a variable related to realized or implied volatility volatility.

In this section, we use the return of the $t = 12$ to $t = 1$ momentum factor from the Jensen-Kelly-Pedersen data depository to conduct the tests. Our findings are reported in Table 2.7. The results in the table are qualitatively identical to those in Table 2.2: momentum delivers positive returns during periods of low volatility and the profitability of the strategy declines significantly during periods of high volatility. The dummy variable specifications show that momentum delivers large negative returns if volatility is within the top 20% (or 10%) of full sample volatility.

2.B.2 Mixed Momentum

In this section, we implement our mixed momentum strategy using the $t = 12$ to $t = 1$ and the $t = 3$ to $t = 1$ momentum factors from the Jensen-Kelly-Pedersen data repository. Figure 2.8 depicts the performance of the restricted and unrestricted versions of the mixed momentum strategy relative to the performance of the baseline momentum strategy.

The difference in performance between the restricted mixed momentum strategy and the baseline momentum strategy is not as stark as the difference in Figure 2.5. If we invested \$1 in the two strategies in 01/1986, the baseline momentum strategy would generate \$2.94 and the restricted mixed strategy would generate \$3.52. However, the mixed strategy works as intended and lessens the extent of momentum crashes during high-volatility episodes. The unrestricted version of the mixed momentum strategy outperforms both the restricted mixed strategy and the baseline strategy by leveraging up and taking on relatively more risk during low-volatility periods.

2.C XGBoost

Formally, the XGBoost algorithm involves minimizing the following objective function for the i -th observation at the t -th iteration:

$$\mathcal{L}^{(t)} = \sum_{i=1}^n \ell \left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i) \right) + \Omega(f_t) \quad (2.51)$$

by greedily adding tree f_t that most improves the model.

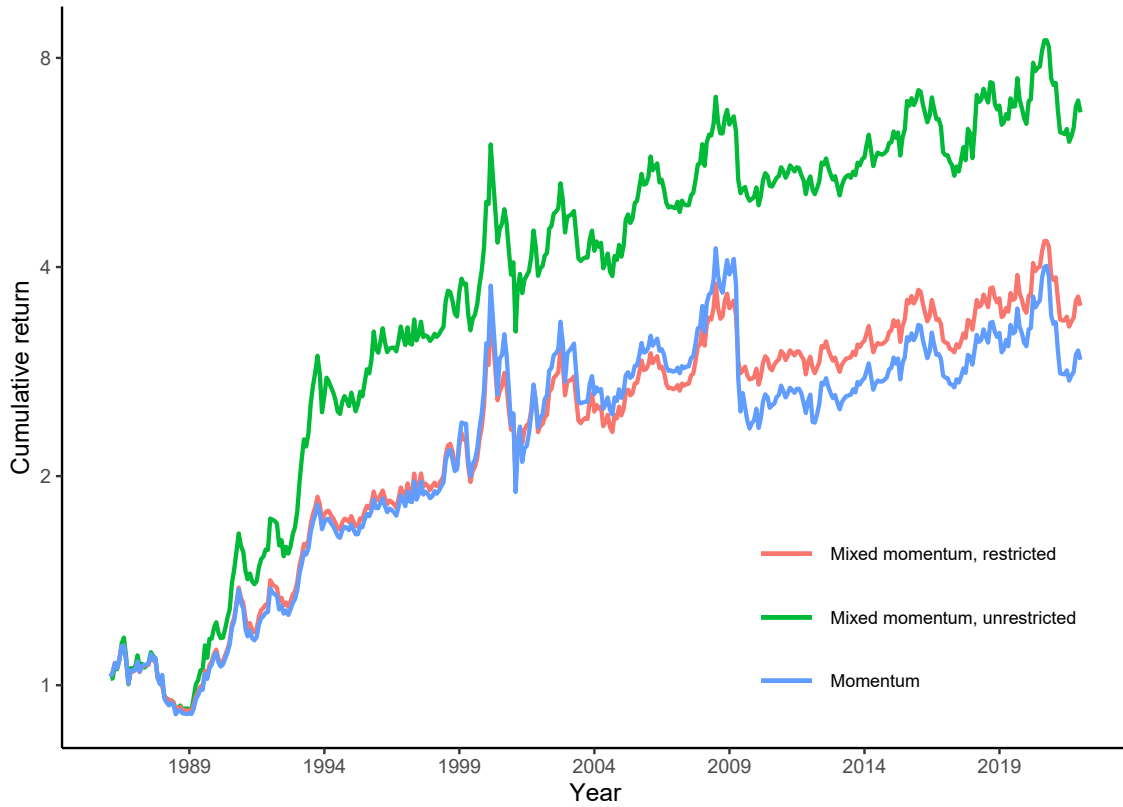
Here $\ell(\cdot)$ represents a loss function that measures the difference between the predicted value of the outcome variable, \hat{y}_i and the realized value y_i , x_i represents the vector of predictors associated with observation i , and the $\Omega(\cdot)$ function penalizes the complexity of the model.

Table 2.7
Momentum and Volatility

	<i>Dependent variable:</i>						
	$r_{\text{WML},t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	RV				IV		
$\hat{\sigma}_{\text{mkt}}$		-1.230* [-1.834]			-0.925** [-2.227]		
$\mathbf{1}_{\text{HV}}$			-0.014* [-1.915]	-0.010 [-0.894]		-0.015** [-2.128]	-0.033*** [-2.935]
Constant	0.004 [1.610]	0.016** [2.578]	0.006*** [3.175]	0.005** [2.017]	0.019*** [3.082]	0.006*** [3.414]	0.007*** [3.318]
Observations	432	432	432	432	422	422	422
$P(\mathbf{1}_{\text{HV}} = 1)$	—	—	0.20	0.10	—	0.20	0.10

This table reports the results for the following regression: $r_{\text{WML},t} = \alpha + \alpha_{\Delta} z_{t-1} + \varepsilon_t$. In columns (2) and (5) z_{t-1} is realized and implied volatility, respectively. In columns (3), (4), (6), and (7) z_{t-1} is a high volatility indicator. The CBOE *S&P100* Volatility Index is used as a measure of implied volatility and the volatility of value-weighted CRSP is used as a measure of realized volatility. The standard errors are computed using the Newey and West 1987 methodology with six lags. The corresponding *t*-statistics are in square brackets. Significance at the 1%, 5%, and 10% is denoted by ***, **, and *, respectively. The momentum return data is obtained from the Jensen-Kelly-Pedersen data repository and covers the period between 01/1986 and 12/2021.

Figure 2.8
Mixed Momentum Strategy



The figure above depicts the cumulative returns of the restricted and unrestricted versions of the mixed momentum trading strategy developed in Section 2.5.2.1, along with the returns of the baseline momentum strategy. We obtain the momentum data from the Jensen-Kelly-Pedersen data repository. We consider the time period between 01/1986 and 12/2021.

References

- Abarbanell, Jeffery S., and Victor L. Bernard. 1992. “Tests of Analysts’ Overreaction/Underreaction to Earnings Information as an Explanation for Anomalous Stock Price Behavior.” *The Journal of Finance* 47 (3): 1181–1207. <https://doi.org/https://doi.org/10.1111/j.1540-6261.1992.tb04010.x>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1992.tb04010.x>.
- Andrade, Philippe, and Hervé Le Bihan. 2013. “Inattentive professional forecasters.” *Journal of Monetary Economics* 60 (8): 967–982. ISSN: 0304-3932. <https://doi.org/https://doi.org/10.1016/j.jmoneco.2013.08.005>.
- Barroso, Pedro, and Pedro Santa-Clara. 2015. “Momentum has its moments.” *Journal of Financial Economics* 116 (1): 111–120. ISSN: 0304-405X. <https://doi.org/https://doi.org/10.1016/j.jfineco.2014.11.010>.
- Barroso, Pedro, and Haoxu Wang. 2022. *What Explains Price Momentum and 52-Week High Momentum When They Really Work?* Working Paper.
- Ben-Rephael, Azi, Zhi Da, and Ryan D. Israelsen. 2017. “It Depends on Where You Search: Institutional Investor Attention and Underreaction to News.” *The Review of Financial Studies* 30, no. 9 (May): 3009–3047. ISSN: 0893-9454. <https://doi.org/10.1093/rfs/hhx031>. eprint: <https://academic.oup.com/rfs/article-pdf/30/9/3009/24434209/hhx031.pdf>.
- Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer. 2022. *Belief Overreaction and Stock Market Puzzles*.

- Bouchaud, Jean-Philippe, Philipp Kruger, Augustin Landier, and David Thesmar. 2019. “Sticky Expectations and the Profitability Anomaly.” *The Journal of Finance* 74 (2): 639–674. <https://doi.org/https://doi.org/10.1111/jofi.12734>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12734>.
- Bryzgalova, Svetlana, Sven Lerner, Martin Lettau, and Sven Pelger. 2022. *Missing Financial Data*.
- Cao, Kai, and Haifeng You. 2021. *Fundamental Analysis via Machine Learning*. Working Paper.
- Chen, Tianqi, and Carlos Guestrin. 2016. “XGBoost: A Scalable Tree Boosting System.” *CoRR* abs/1603.02754. arXiv: 1603.02754. <http://arxiv.org/abs/1603.02754>.
- Chordia, Tarun, Avanidhar Subrahmanyam, and Qing Tong. 2014. “Have capital market anomalies attenuated in the recent era of high liquidity and trading activity?” *Journal of Accounting and Economics* 58 (1): 41–58. ISSN: 0165-4101. <https://doi.org/https://doi.org/10.1016/j.jacceco.2014.06.001>.
- Coibion, Olivier, and Yuriy Gorodnichenko. 2015. “Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts.” *American Economic Review* 105 (8): 2644–78.
- Da, Zhi, Umit G. Gurun, and Mitch Warachka. 2014. “Frog in the Pan: Continuous Information and Momentum.” *The Review of Financial Studies* 27, no. 7 (February): 2171–2218. ISSN: 0893-9454. <https://doi.org/10.1093/rfs/hhu003>. eprint: <https://academic.oup.com/rfs/article-pdf/27/7/2171/24449796/hhu003.pdf>.

- Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam. 1998. “Investor Psychology and Security Market under- and Overreactions.” *The Journal of Finance* 53 (6): 1839–1885. ISSN: 00221082, 15406261, accessed August 9, 2022. <http://www.jstor.org/stable/117455>.
- Daniel, Kent, and Tobias J. Moskowitz. 2016. “Momentum crashes.” *Journal of Financial Economics* 122 (2): 221–247. ISSN: 0304-405X. <https://doi.org/https://doi.org/10.1016/j.jfineco.2015.12.002>.
- Dellavigna, Stefano, and Joshua M. Pollet. 2009. “Investor Inattention and Friday Earnings Announcements.” *The Journal of Finance* 64 (2): 709–749. <https://doi.org/https://doi.org/10.1111/j.1540-6261.2009.01447.x>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2009.01447.x>.
- Diether, Karl B., Christopher J. Malloy, and Anna Scherbina. 2002. “Differences of Opinion and the Cross Section of Stock Returns.” *The Journal of Finance* 57 (5): 2113–2141. ISSN: 00221082, 15406261, accessed September 22, 2022. <http://www.jstor.org/stable/3094506>.
- Engelberg, Joseph, R. David McLean, and Jeffrey Pontiff. 2018. “Anomalies and News.” *The Journal of Finance* 73 (5): 1971–2001. <https://doi.org/https://doi.org/10.1111/jofi.12718>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12718>.
- Fama, Eugene F., and Kenneth R. French. 2015. “A five-factor asset pricing model.” *Journal of Financial Economics* 116 (1): 1–22. ISSN: 0304-405X. <https://doi.org/https://doi.org/10.1016/j.jfineco.2014.10.010>.

- Frankel, Richard, and Charles M.C. Lee. 1998. “Accounting valuation, market expectation, and cross-sectional stock returns.” *Journal of Accounting and Economics* 25 (3): 283–319. ISSN: 0165-4101. [https://doi.org/https://doi.org/10.1016/S0165-4101\(98\)00026-3](https://doi.org/https://doi.org/10.1016/S0165-4101(98)00026-3).
- Goulding, Christian L., Campbell R. Harvey, and Michele Mazzoleni. 2022. *Momentum Turning Points*. Working Paper.
- Harvey, Campbell R., Yan Liu, and Heqing Zhu. 2015. “ and the Cross-Section of Expected Returns.” *The Review of Financial Studies* 29, no. 1 (October): 5–68. ISSN: 0893-9454. <https://doi.org/10.1093/rfs/hhv059>. eprint: <https://academic.oup.com/rfs/article-pdf/29/1/5/24450794/hhv059.pdf>.
- Hirshleifer, David, Sonya Seongyeon Lim, and Siew Hong Teoh. 2009. “Driven to Distraction: Extraneous Events and Underreaction to Earnings News.” *The Journal of Finance* 64 (5): 2289–2325. <https://doi.org/https://doi.org/10.1111/j.1540-6261.2009.01501.x>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2009.01501.x>.
- Hong, Harrison, Terence Lim, and Jeremy C. Stein. 2000. “Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies.” *The Journal of Finance* 55 (1): 265–295. <https://doi.org/https://doi.org/10.1111/0022-1082.00206>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/0022-1082.00206>.

- Hong, Harrison, and Jeremy C. Stein. 1999. “A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets.” *The Journal of Finance* 54 (6): 2143–2184. ISSN: 00221082, 15406261, accessed August 2, 2022. <http://www.jstor.org/stable/797990>.
- Hou, Kewei, Chen Xue, and Lu Zhang. 2014. “Digesting Anomalies: An Investment Approach.” *The Review of Financial Studies* 28, no. 3 (September): 650–705. ISSN: 0893-9454. <https://doi.org/10.1093/rfs/hhu068>. eprint: <https://academic.oup.com/rfs/article-pdf/28/3/650/24450149/hhu068.pdf>.
- Hribar, Paul, and Nicole Thorne Jenking. 2004. “The Effect of Accounting Restatements on Earnings Revisions and the Estimated Cost of Capital.” *Review of Accounting Studies* 9:337356.
- Jegadeesh, Narasimhan, and Sheridan Titman. 1993. “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency.” *The Journal of Finance* 48 (1): 65–91. ISSN: 00221082, 15406261, accessed September 26, 2022. <http://www.jstor.org/stable/2328882>.
- Jensen, Theis Ingerslev, Bryan T Kelly, and Lasse Heje Pedersen. 2022. “Is There A Replication Crisis In Finance?” *Journal of Finance, Forthcoming*.
- Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp. 2016. “A Rational Theory of Mutual Funds’ Attention Allocation.” *Econometrica* 84 (2): 571–626. <https://doi.org/https://doi.org/10.3982/ECTA11412>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA11412>.

- La Porta, Rafael. 1996. "Expectations and the Cross-Section of Stock Returns." *The Journal of Finance* 51 (5): 1715–1742. ISSN: 00221082, 15406261, accessed October 23, 2022. <http://www.jstor.org/stable/2329535>.
- Loungani, Prakash, Herman Stekler, and Natalia Tamirisa. 2013. "Information rigidity in growth forecasts: Some cross-country evidence." *International Journal of Forecasting* 29 (4): 605–621. ISSN: 0169-2070. <https://doi.org/https://doi.org/10.1016/j.ijforecast.2013.02.006>.
- Luo, Jiang, Avanidhar Subrahmanyam, and Sheridan Titman. 2020. "Momentum and Reversals When Overconfident Investors Underestimate Their Competition." *The Review of Financial Studies* 34, no. 1 (February): 351–393. ISSN: 0893-9454. <https://doi.org/10.1093/rfs/hhaa016>. eprint: <https://academic.oup.com/rfs/article-pdf/34/1/351/34998838/hhaa016.pdf>.
- Mankiw, N. Gregory, and Ricardo Reis. 2002. "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *The Quarterly Journal of Economics* 117, no. 4 (November): 1295–1328. ISSN: 0033-5533. <https://doi.org/10.1162/003355302320935034>. eprint: <https://academic.oup.com/qje/article-pdf/117/4/1295/5304341/117-4-1295.pdf>.
- Martineau, Charles. 2023. "Rest in Peace Post-Earnings Announcement Drift." *Critical Finance Review*.
- Moreira, Alan, and Tyler Muir. 2017. "Volatility-Managed Portfolios." *The Journal of Finance* 72 (4): 1611–1644. <https://doi.org/https://doi.org/10.1111/jofi.12513>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12513>.

- Newey, Whitney K., and Kenneth D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55 (3): 703–708. ISSN: 00129682, 14680262, accessed August 2, 2022. <http://www.jstor.org/stable/1913610>.
- Novy-Marx, Robert. 2013. "The other side of value: The gross profitability premium." *Journal of Financial Economics* 108 (1): 1–28. ISSN: 0304-405X. <https://doi.org/https://doi.org/10.1016/j.jfineco.2013.01.003>.
- Palmrose, Zoe-Vonna, Vernon J. Richardson, and Susan Scholz. 2004. "Determinants of market reactions to restatement announcements." *Journal of Accounting and Economics* 37 (1): 59–89.
- Pouget, Sebastien, Julien Sauvagnat, and Stephane Villeneuve. 2016. "A Mind Is a Terrible Thing to Change: Confirmatory Bias in Financial Markets." *The Review of Financial Studies* 30, no. 6 (December): 2066–2109. ISSN: 0893-9454. <https://doi.org/10.1093/rfs/hhw100>. eprint: <https://academic.oup.com/rfs/article-pdf/30/6/2066/24434108/hhw100.pdf>.
- Silva, Tim de, and David Thesmar. 2023. *Noise in Expectations: Evidence from Analyst Forecasts*. Working Paper.
- So, Eric C. 2013. "A new approach to predicting analyst forecast errors: Do investors overweight analyst forecasts?" *Journal of Financial Economics* 108 (3): 615–640. ISSN: 0304-405X. <https://doi.org/https://doi.org/10.1016/j.jfineco.2013.02.002>.

Van Binsbergen, Jules H, Xiao Han, and Alejandro Lopez-Lira. 2022. “Man versus Machine Learning: The Term Structure of Earnings Expectations and Conditional Biases.” *The Review of Financial Studies* (October). ISSN: 0893-9454. <https://doi.org/10.1093/rfs/hhac085>. eprint: <https://academic.oup.com/rfs/advance-article-pdf/doi/10.1093/rfs/hhac085/46698800/hhac085.pdf>.

White, Halbert. 1980. “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” *Econometrica* 48 (4): 817–838. ISSN: 00129682, 14680262, accessed September 13, 2022. <http://www.jstor.org/stable/1912934>.