Escola de Pós-Graduação em Economia - EPGE
Fundação Getúlio Vargas

Essays on Regulation and Risk

Tese submetida à Escola de Pós-Graduação em Economia
da Fundação Getúlio Vargas como requisito de obtenção do
título de Doutor em Economia

Aluno: Régio Soares Ferreira Martins

Orientador: Humberto Luiz Ataíde Moreira

Rio de Janeiro
2010

Escola de Pós-Graduação em Economia - EPGE
Fundação Getúlio Vargas

Essays on Regulation and Risk

Tese submetida à Escola de Pós-Graduação em Economia
da Fundação Getúlio Vargas como requisito de obtenção do
título de Doutor em Economia

Aluno: Régio Soares Ferreira Martins

Banca Examinadora:
Humberto Luiz Ataíde Moreira (Orientador, EPGE-FGV)
Renato Galvão Flôres Junior (EPGE/FGV)
Caio Ibsen Rodrigues de Almeida (EPGE/FGV)
Vinícius Carrasco (PUC/RJ)
Rodrigo de Losso da Silveira Bueno (FEA/USP)

Rio de Janeiro
2010

Dedico esse trabalho aos meus pais, Marcos e Helena, pela educação que deles recebi; e à Daniella, pela paciência que sempre teve comigo.

Agradeço,

Ao Humberto, pelas discussões, preocupações e cobranças que foram fundamentais para a conclusão desse trabalho.
Aos membros da banca examinadora, pela paciência na leitura dos artigos, sugestões e constribuições feitas.
Ao meu co-autor, Daniel Lima, pelas ideias, discussões (e diversões) em San Diego.
Aos colegas da EPGE, pela experiência que tivemos e pela amizade que ficou.
Ao Edson Gonçalves, cuja ajuda foi fundamental para o término do trabalho.
Ao CNpQ e à FGV, pelo suporte financeiro e intelectual.

RESUMO

Essa tese investiga alguns aspectos da relação entre regulação econômica e risco da empresa regulada. No primeiro capítulo, o objetivo é entender as implicações do modelo tradicional de regulação por incentivos (Laffont e Tirole, 1993) sobre o risco sistemático da firma. Generalizamos o modelo de forma a incorporar risco agregado ao lucro da atividade, e descobrimos que o contrato ótimo deve ser severamente restringido para que reproduza betas (CAPM) próximos aos observados em setores regulados. Usamos um caso particular do modelo, de regulação por repartição de lucro (*profit-sharing regulation*), para avaliar a relação entre a potência do contrato e o nível de risco não diversificável. Encontramos resultados compatíveis com a evidência disponível, de que regimes com alta potência impõem mais risco sobre a firma.

No segundo capítulo, escrito em co-autoria com Daniel Lima da Universidade da Califórnia em San Diego (UCSD), partimos da constatação de que empresas reguladas podem estar sujeitas a práticas regulatórias que potencialmente afetam a simetria da distribuição de seus lucros futuros. Se essas práticas forem antecipadas pelos investidores no mercado secundário de ações, poderemos identificar diferenças no padrão da assimetria da distribuição empírica de retornos das empresas reguladas com relação às não-reguladas. Nesse capítulo revisamos alguns métodos de mensuração de assimetria propostos recentemente na literatura, que são robustos à características comuns em séries de retornos financeiros (caudas pesadas e correlação serial), e investigamos se existem diferenças significativas na distribuição de assimetria entre empresas reguladas e não-reguladas.

No terceiro e último capítulo, três diferentes abordagens empíricas do modelo de apreçamento de ativos de Kraus e Litzenberger (1976) são testadas com dados do mercado brasileiro de ações. Descobrimos que a distribuição empírica de retornos costuma exibir co-assimetria significativa com relação à carteira de mercado, e que portanto os retornos das ações são sensíveis à volatilidade (retornos quadráticos) do mercado. No entanto, apesar da base teórica para a preferência por retornos assimétricos esteja bem estabelecida e seja bastante intuitiva, não encontramos evidência que suporte a hipótese de que os investidores requeiram um prêmio para aceitar esse tipo de risco no mercado local.

Palavras-chave: Regulação por incentivos, CAPM, risco regulatório, assimetria, co-assimetria, métodos não-paramétricos.

ABSTRACT

In this thesis, we investigate some aspects of the interplay between economic regulation and the risk of the regulated firm. In the first chapter, the main goal is to understand the implications a mainstream regulatory model (Laffont and Tirole, 1993) have on the systematic risk of the firm. We generalize the model in order to incorporate aggregate risk, and find that the optimal regulatory contract must be severely constrained in order to reproduce real-world systematic risk levels. We also consider the optimal profit-sharing mechanism, with an endogenous sharing rate, to explore the relationship between contract power and beta. We find results compatible with the available evidence that high-powered regimes impose more risk to the firm.

In the second chapter, a joint work with Daniel Lima from the University of California, San Diego (UCSD), we start from the observation that regulated firms are subject to some regulatory practices that potentially affect the symmetry of the distribution of their future profits. If these practices are anticipated by investors in the stock market, the pattern of asymmetry in the empirical distribution of stock returns may differ among regulated and non-regulated companies. We review some recently proposed asymmetry measures that are robust to the empirical regularities of return data and use them to investigate whether there are meaningful differences in the distribution of asymmetry between these two groups of companies.

In the third and last chapter, three different approaches to the capital asset pricing model of Kraus and Litzenberger (1976) are tested with recent Brazilian data and estimated using the generalized method of moments (GMM) as a unifying procedure. We find that ex-post stock returns generally exhibit statistically significant coskewness with the market portfolio, and hence are sensitive to squared market returns. However, while the theoretical ground for the preference for skewness is well established and fairly intuitive, we did not find supporting evidence that investors require a premium for supporting this risk factor in Brazil.

Keywords: Incentive regulation, CAPM, regulatory risk, skewness, coskewness, non-parametric methods.

Contents

List of Tables

List of Figures

Chapter 1

Incentive Regulation and Beta

We generalize the Laffont and Tirole (1993) model of regulation in order to incorporate aggregate risk in the firm's profit. We find that the optimal regulatory contract must be severely constrained in order to reproduce real-world systematic risk levels. We also consider the optimal profit-sharing mechanism, with an endogenous sharing rate, to explore the relationship between contract power and beta. We find results compatible with the available evidence that high-powered regimes impose more risk to the firm.

Keywords: Incentive regulation, CAPM, beta, regulatory risk.

JEL Classification: L51, L140

1.1 Introduction

Regulatory agencies are the primary forum where interest groups exert their influence and lobby for a favorable policy. As a consequence of the political process, regulators usually have many different goals to pursue and instruments at their disposal. The modern economic theory of regulation[1] works with a more parsimonious characterization of the institutional framework and objectives of regulators, assuming that they choose a few instruments such as retail prices, quality standards, rate revision rules or legal barriers to entry and exit towards maximizing some concept of social welfare. But given the complex nature of their business and environment, regulators usually keep a certain amount of discretion in conducting policy as complete contracts are impractical to be written.

When setting output prices and the rules for their revision, the regulator inevitably allocates risk between consumers and the firm. The form of price regulation has a major role in determining which group will support most of demand or supply shocks. In low-powered regimes such as the cost-plus, prices are frequently adjusted so that the firm is guaranteed an explicit rate of return on its invested capital, and hence consumers bear most cost and demand fluctuations. But in high-powered systems such as the price-cap, prices cannot be raised above the predetermined ceiling between rate hearings, so the firm is the residual claimant of any cost hikes or reductions and must support some of the demand risk as well.

Usually the regulated service is supplied by one or few firms, so the price rule must strike a balance between limiting market power and allowing the firm to recover the investment made. In the longer term prices should be cost-oriented, and the required rate of return on the firm's capital is an important part of the total cost to be covered[2]. As Guthrie (2006) explains it, the allowed price for period $t+1$ should generate an expected revenue that satisfies $\mathbb{E}_t[R_{t+1}] = \mathbb{E}_t[C_{t+1}] + r_t B_t + D_{t+1}$, where C_{t+1} is the operating cost, r_t is the opportunity cost of capital, B_t is the invested capital or rate

[1] Armstrong and Sappington (2007) provide a thorough survey of the recent developments in the field.

[2] This is the generally accepted view at least since the Federal Power Commission vs. Hope Natural Gas & Co. case in 1944, when the U.S. Supreme Court decided that "the return to the equity owner should be commensurate with returns on investments in other enterprises having corresponding risks. That return, moreover, should be sufficient to assure confidence in the financial integrity of the enterprise so as to maintain its credit and to attract capital". Myers (1972) explains how finance theory, and the CAPM in particular, can be used to implement these principles, while Wright, Mason and Miles (2003) and Jenkinson (2006) provide a recent view on how regulators should measure the cost of capital.

base and D_{t+1} is the depreciation. If the regulator calculates depreciation as $D_t = B_t - \mathbb{E}_t[B_{t+1} - I_{t+1}]$, where I_{t+1} is the next period's investment, and allows a rate of return equal to the opportunity cost of capital, then the market value of the firm $V_t = \mathbb{E}_t[R_{t+1} - C_{t+1} - I_{t+1} + V_{t+1}]/(1 + r_t)$ will equal the rate base B_t. However, the opportunity cost of capital is not exogenous but partially determined by the risk allocation embedded in the regulatory policy. Marshall, Yawitz and Greenberg (1981) and Brennan and Schwartz (1982) were the first attempts to model the simultaneous determination of output prices and rates of return under uncertainty. Brennan and Schwartz call *consistent regulation* one in which allowed prices are compatible with $V_t = B_t$ at every t. These models, however, were developed under the cost-plus benchmark, before the advent of the modern incentive theory of regulation. In this paper we explicitly account for the joint determination of equilibrium output prices and rates of return under optimal incentive schemes, so our model yield consistent policies in the sense of Brennan and Schwartz and consider productive incentives at the same time.

It is not entirely clear, however, whether incentive regulation increases the risk of the firm when compared to the traditional rate of return system. For instance, firms under a price-cap have downward price flexibility, which mitigates negative demand shocks and lowers the cost of capital. Also, their rate base and allowed prices are revised less often, diminishing the additional uncertainty brought by the regulatory process. On the other hand, they face more risk concerning the valuation of the rate base, which in many cases follows the efficient (or replacement) principle. Due to all these opposing effects, the end result is fundamentally an empirical issue. There is some evidence that higher-powered systems raise the firm's CAPM beta. Alexander, Mayer and Weeds (1996) compare a cross-section of regulated firms from several countries and industries and find higher average betas where regulation is more powered, while Alexander, Estache and Oliveri (2000) find the same evidence for the global transportation sector. Grout and Zalewska (2006) show how betas of regulated firms in the U.K. decreased during the period in which the Labour Party tried to replace the prevailing price-cap system with a profit-sharing one.

However, empirical studies of the relationship between regulatory incentives and risk are fraught with difficulties. It is extremely difficult to find exogenous variation in contract power for a significant number of regulated firms and to control for all relevant confounding factors, which makes cross-sectional approaches particularly unfruitful. The next alternative is to focus on a single industry with many regulated entities that have been subject to changes in the degree of incentives. Schwert (1981) and Binder (1985) provide some methodology for event studies of regulatory changes but emphasize

three major challenges in such studies: (i) the lack of a clear event window where an announcement is made by the authority and market expectations change; (ii) the non-uniform impact of the new rules across regulated firms; and (iii) the fact that an announcement may affect many companies in the same industry at the same time, and may be confounded with exogenous industry-specific shocks.

It is therefore clear that the interaction between regulation and the firm's cost of capital is not only important but complex in nature and hard to be measured empirically. It is useful to understand how the regulatory system creates systematic risk, specially under a consistent policy with simultaneous determination of retail prices and betas. Even if the risk characteristics of extreme forms of price regulation such as the cost-plus and price-cap could be understood, Laffont and Tirole (1986, 1993) showed that the optimal regulatory contract has an intermediate power between these two extremes, depending on the cost efficiency of the firm. It is therefore of natural interest to investigate how risk and the power of the regulatory contract interact.

The regulatory model developed in this article starts from the approach of Laffont and Tirole. Then, we allow for aggregate uncertainty in a non-trivial way, as the overall performance of the economy exerts some influence on the cost efficiency of the regulated firm and partially influences policy. Since the regulatory contract is a set of state-contingent production plans and rents, the link between the firm's type and the aggregate shocks creates systematic risk in the regulated business. As a consequence, the firm's rational entry decision is based on a risk-adjusted valuation of the uncertain profits, and the regulator must explicitly account for this when designing the contract terms, a feature that is also absent in the original Laffont and Tirole model.

1.2 Model

We consider a one-period economy with complete financial markets and an aggregate market portfolio whose return is denoted by r_m. Without loss of generality, this return is decomposed as $r_m = r_f + \gamma + \tilde{m}$, where $r_f > 0$ is the risk-free rate, $\tilde{m} \in \{m_L, m_H\}$ is a random shock with zero mean and finite variance σ_m^2 and $\gamma \equiv \mathbb{E}[r_m - r_f] > 0$ is the market risk premium.

There is a single-product firm with access to a constant returns to scale technology represented by the cost function $C(q, e, \tilde{\theta}) = (\tilde{\theta} - e)q$, where $\tilde{\theta} \in \{\theta_L, \theta_H\}$ is an exogenous and random efficiency parameter and $e \geq 0$ is the level of effort chosen by the manager-shareholder to reduce costs. The firm's profit or cash flow is $U = t - \Psi(e)$, where t is a transfer (subsidy) received from the regulator and $\Psi(e) = \psi\, e^2/2$, $\psi > 0$, is the "disutility"

from exerting effort, measured in monetary terms as a cost incurred by the firm's manager. The industry is small relative to the overall economy so the impact of the firm's earnings on the performance of the market portfolio is negligible.

Consumers have inverse demand $p(q, \tilde{m}) = d(1 + \nu\,\tilde{m}) - q$ for the good, where $d > 0$ is a taste parameter and $0 < \nu < 1/\sigma_m$ reflects how sensitive demand is to the aggregate shock $\tilde{m}$. The gross surplus derived from the consumption of q units is $S(q, \tilde{m}) \equiv \int_0^q p(z, \tilde{m})dz$.

The random shocks $\tilde{\theta}$ and $\tilde{m}$ can assume either a high or low value, so we characterize a state of nature by a realization $(\theta_j, m_k) \in \{\mu_\theta - \sigma_\theta, \mu_\theta + \sigma_\theta\} \times \{-\sigma_m, \sigma_m\}$, where $\sigma_\theta > 0$ is the standard deviation and $\mu_\theta > \sigma_\theta$ is the expected value of $\tilde{\theta}$. Throughout this paper, we use $j = L, H$ to index realizations of $\tilde{\theta}$ and $k = L, H$ for realizations of $\tilde{m}$.

The firm's technology is partially determined by the aggregate performance of the economy, and we let ρ denote the correlation between $\tilde{\theta}$ and $\tilde{m}$. The probability of state (θ_j, m_k) is $(1+\rho)/4$ if $j = k$, and $(1-\rho)/4$ otherwise. This joint probability distribution implies that the (unconditional) marginal probability of each shock being either low or high is the same, $\sum_k P(\theta_j, m_k) = 1/2$ for any j and $\sum_j P(\theta_j, m_k) = 1/2$ for any k.

We follow the general framework of the Laffont and Tirole (1986, 1993) model of regulation under adverse selection and moral hazard, in which the realized cost and output are observable, but effort and efficiency are not. In this model, the regulator reimburses the realized production cost $C(q, e, \theta_j)$ and pays a net subsidy t to the firm using public funds with marginal cost $\lambda > 0$. The regulator is utilitarian and maximizes the sum of net surplus with firm's cash flow, $S(q, m_k) - (1+\lambda)[C(q, e, \theta_j) + t] + U$, using (C, q, t) as policy instruments. The *ex-post* social surplus function can be re-written in the (q, e, U)-space as

$$W_{jk}(q, e, U) \equiv S(q, m_k) - (1+\lambda)[(\theta_j - e)q + \Psi(e)] - \lambda U, \tag{1.1}$$

to become explicit that rent is socially costly. As $\Psi(\cdot)$ is strictly increasing and $\tilde{\theta}$ can be observed (or correctly inferred through a direct revelation mechanism), there is a one-to-one mapping from contracts in the (C, q, t)-space to the (q, e, U)-space. Welfare is globally concave in (q, e) if and only if

$$\psi > 1 + \lambda, \tag{1.2}$$

so the first order conditions for the maximization of (1.1) are sufficient for a global optimum when the marginal disutility of effort is bounded below by $(1 + \lambda)$.

Let (q_{jk}, e_{jk}, U_{jk}) denote the output, effort and utility of the firm in state (θ_j, m_k). The firm operates under limited liability, and we normalize her *ex-post* cash flow to be non-negative in every state,

$$U_{jk} \geq 0 \quad j, k = L, H. \tag{1.3}$$

Let $I \geq 0$ be the required (and verifiable) capacity level. Then

$$q_{jk} \leq I \quad j, k = L, H \tag{1.4}$$

is the feasibility constraint. There is no capacity in place in the beginning of the period, and the full investment cost of $\$I$ should be borne by the firm. Capacity takes time to be built, so it must be done *ex-ante* and cannot be made contingent on the state. For simplicity, we also assume that this investment is completely sunk and fully depreciates at the end of the period.

A regulatory policy specifies a capacity level and a set of state-contingent production plans and rents

$$\{I, (q_{jk}, e_{jk}, U_{jk})_{j,k=L,H}\} \tag{1.5}$$

that is feasible and satisfies limited liability, i.e., (1.3) and (1.4) hold. The regulatory policy must be revealed to the firm in the beginning of the period, before uncertainty is resolved, so that she may decide whether or not to enter the regulated business. The firm's entry decision is modeled as follows. Let $r_{jk} = U_{jk}/V - 1$ denote the net return in state (θ_j, m_k), where V is the value of the firm's equity at the beginning of the period. The expected return on the regulated business is $\mathbb{E}[r] = \mathbb{E}[U]/V - 1$. In a CAPM equilibrium the required rate of return on the firm's equity is $\mathbb{E}[r] = r_f + \beta\,\mathbb{E}[r_m - r_f]$, thus

$$\frac{\mathbb{E}[U]}{V} - 1 = r_f + \frac{\operatorname{cov}(r, r_m)}{\sigma_m^2}\mathbb{E}[r_m - r_f] = r_f + \frac{\mathbb{E}[Um]}{V\,\sigma_m^2}\gamma,$$

where the second equality follows from the fact that $\tilde{m}$ is a zero mean shock. Solving the above for V yields the risk-neutral (or certainty equivalent) formula for V. The firm shareholders will invest if and only if

$$V = \frac{\mathbb{E}[U] - (\gamma/\sigma_m^2)\mathbb{E}[Um]}{1 + r_f} \geq I. \tag{1.6}$$

Consequently, the systematic risk β implicit in the regulated business is given by

$$\beta = \frac{\mathbb{E}[Um](1 + r_f)}{\sigma_m^2\mathbb{E}[U] - \gamma\,\mathbb{E}[Um]}. \tag{1.7}$$

The denominator in (1.7) is simply the value of the firm multiplied by the variance of r_m, so it must be positive in equilibrium (the firm would not be created if $V < 0$), and the sign of β will be determined by the sign of the covariance between the firm's cash flows and the market shocks, $\mathbb{E}[Um]$. Also, one can also show that $\partial\beta/\partial U_{jk} > 0$ if and only if $k = H$, so allocating rent in good states increases β.

The regulator's problem is to choose a policy (1.5) that maximizes expected welfare

$$\mathbb{E}[W] \equiv \sum_j \sum_k P(\theta_j, m_k) W_{jk}(q_{jk}, e_{jk}, U_{jk}) \tag{1.8}$$

where $j, k = L, H$, subject to (1.3), (1.4), (1.6) and possibly other restrictions imposed by the contracting environment, such as the impossibility of transfers or incomplete information about the firm's type or realized cost. We will characterize the optimal set of contracts under full commitment in Section 1.3.

1.3 Regulatory Policies

When the regulator is able to credibly commit to a policy in which (1.6) also holds, there is no problem of inducing investment at the beginning of the period. The firm can be certain to receive a rent of U_{jk} when the state is (θ_j, m_k) and consequently is valued at least as much as the initial investment. Thus, by assumption, the firm sunks capital before uncertainty is resolved as there is no default risk from the regulator.

1.3.1 Complete Information (FB)

First, we characterize the first-best output and effort levels under no capacity or financial restrictions, obtained by the unconstrained maximization of (1.8):

$$e^*_{jk} = \frac{d(1 + \nu\, m_k) - (1 + \lambda)\theta_j}{\psi - (1 + \lambda)}, \tag{1.9a}$$

$$q^*_{jk} = \psi\, e^*_{jk}. \tag{1.9b}$$

When $d\,\nu\,\sigma_m \geq (1 + \lambda)\sigma_\theta$, we have $q^*_{LH} > q^*_{LL} \geq q^*_{HH} > q^*_{HL}$, otherwise $q^*_{LH} > q^*_{HH} > q^*_{LL} > q^*_{HL}$. In any case, the maximal desired output is implemented when cost is low and demand is high. In order to simplify the analysis, from now on we assume a fixed capacity at this maximal level,

$$I = q^*_{LH}, \tag{1.10}$$

and refrain from discussing its optimal determination under the different policies that we study. Obviously, the regulator can choose the capacity optimally and simultaneously with the other variables, but we avoid to do so as this would complicate the analysis unnecessarily. An exogenously fixed investment level is sufficient to raise the problem of rent allocation across states and the issues associated with the absence of commitment. In Appendix A.1, we show that the optimal capacity level under complete information is distorted downwards from q^*_{LH} by an amount that reflects the social cost of funds, as the full cost of the investment must be borne by the firm which is indirectly financed by taxpayers. For instance, when the marginal cost of funds λ or aggregate uncertainty σ_m are low, (1.10) is a good approximation of the optimal capacity in the first-best.

The timing of regulation is the following. First, the regulator designs a set of state-contingent contracts (1.5) conditional on the firm to be willing to participate, i.e. under (1.3) and (1.6). The firm observes this set of contracts and starts building the capacity I in (1.10), which effectively slacks the constraints (1.4) so they can be ignored. Then, a state of nature (θ_j, m_k) realizes and is observed by all agents, so the firm executes the production plan $\{e_{jk}, q_{jk}\}$ and collects a rent equal to U_{jk}.

Under the complete information case, the maximization of (1.8) subject to (1.3) and (1.6) yields the same effort and output given by the unrestricted solution (1.9a) and (1.9b) above. The (implicit) price that would prevail if customers paid for the good directly is

$$p^*_{jk} = (\theta_j - e^*_{jk}) + \frac{\psi}{\psi - (1+\lambda)}\lambda \left[\theta_j - \frac{d(1+\nu m_k)}{\psi}\right], \tag{1.11}$$

so it is optimal to set a state-contingent mark-up that is counter-cyclical (decreases when $\tilde{m} = m_H$) and proportional to the shadow cost of public funds. Also, $\lim_{\psi\to\infty} p^*_{jk} = (1+\lambda)\theta_j$, so when the cost of inducing effort is too high, price is set to the basic marginal (social) cost.

The optimal payment profile related with (1.9) and (1.10) is a set of non-negative, contingent rents that minimizes its expected value $\mathbb{E}[U]$ subject to a binding financial restriction (1.6), and is given by

$$U^*_{LH} = U^*_{HH} = 0, \tag{1.12a}$$

$$U^*_{LL} = \frac{4I(1+r_f)}{(1+\rho)(1+\gamma/\sigma_m)} - \frac{1-\rho}{1+\rho}U^*_{HL}, \tag{1.12b}$$

see Appendix A.1.1 for a derivation. The optimal profile above requires no rent in “good” states, when $\tilde{m} = m_H$, and any combination of non-negative

values of U_{LL} and U_{HL} that satisfy (1.12b). The expected cost of (1.12) is

$$\mathbb{E}[U^*] = I\frac{1+r_f}{1+\gamma/\sigma_m}, \tag{1.13}$$

which is less than I if we replace the parameters in (1.13) by estimates based on their historical averages. If we interpret time in our one-period economy as how long it takes for an entire investment cycle to mature in an infrastructure sector (for instance, five years) and if we base our estimates on the U.S. capital markets, we can use the returns on short-term Treasury bonds and on the S&P 500 as proxies for the risk-free rate and market return. Historical data from the Stocks, Bonds, Bills and Inflation (SBBI) Yearbooks for the period from 1925 to 2005 and averaged for five-year holding periods yield $r_f = 31\%$, $\gamma = 34\%$ and $\sigma_m = 54\%$ (see Table A.2 in Section 1.4). These figures indicate that the expected value of the firm's compensation would be approximately 80% of the investment cost.

The regulator achieves such a low cost of inducing investment by turning the firm into a hedge asset, as (1.12) is a counter-cyclical compensation profile. In effect, under (1.12) the firm's beta is negative:

$$\beta_{FB} = -\frac{1+r_f}{\gamma+\sigma_m}, \tag{1.14}$$

where β_{FB} denotes the firm's beta under the first-best payment scheme. The level of risk attained in (1.14) can be considered the lowest bound for the regulated β, and is the result of a very powerful regulator, who has perfect information about the firm's operating environment and can freely allocate transfers across states and directly to the firm. The historical values for the risk-free rate and market return mentioned before would result in $\beta_{FB} \approx -1.50$, which is unrealistic as real betas are rarely negative: estimates provided by A. Damodaran for global industry betas in 2009 range from 0.03 to 2.73. Table A.1 shows average (unlevered) betas for traditionally regulated services, which range from 0.24 to 0.55.

The regulation specified in (1.9a), (1.9b), (1.10) and (1.12) imposes some basic parameter restrictions for equilibrium values to be sensible. First, effort must be positive, and we have $e^*_{jk} > 0$ for all j, k when the good is socially valuable enough,

$$d > \frac{(1+\lambda)\theta_H}{1-\nu\,\sigma_m}. \tag{1.15}$$

Second, marginal cost must also be positive, i.e., $e^*_{jk} < \theta_j$ for all j, k which requires the marginal disutility of effort to be high enough,

$$\psi > \frac{d(1+\nu\,\sigma_m)}{\theta_L}. \tag{1.16}$$

It is easy to see that (1.15) and (1.16) imply (1.2), so these two are the minimum set of conditions on the exogenous parameters that guarantee concavity of the welfare function and a well-behaved solution under complete information and full commitment. From now on we assume both (1.15) and (1.16) hold, so for each choice of the parameters $\{\mu_\theta, \sigma_\theta, \sigma_m, \gamma, \nu, r_f, \rho, \lambda\}$ that obey the basic primitive restrictions discussed in Section 1.2 (and summarized here for convenience: $\mu_\theta > \sigma_\theta > 0$, $\sigma_m > 0$, $\gamma > 0$, $0 < \nu < 1/\sigma_m$, $r_f > 0$, $\rho \in (-1, 1)$ and $\lambda > 0$), the set of feasible economies under complete information is denoted by $\mathcal{E}^* = \{(d, \psi) \in \mathbb{R}^2_+ \mid \text{(1.15) and (1.16) are valid}\}$.

1.3.2 Asymmetric Information

Now consider the case in which, after entry and sunking investment, the firm privately observes its efficiency parameter θ_j and effort level e. In this case, the regulator observes only the total cost C and output q and cannot disentangle the different components of marginal cost $C/q = \theta - e$. The regulator also observes the market return m_k and could, theoretically, use this information to completely eliminate the firm's advantage. Riordan and Sappington (1988) and Crémer and McLean (1988) showed that, under some assumptions, the principal can recover the first-best outcome by conditioning the contracts on any *ex-post* verifiable signal that conveys some (whatever small) information about the agent's type. However, as pointed out by Demougin and Garvie (1991), this result is not generally applicable in regulatory environments as it requires a strong assumption: the possibility of unbounded transfers (of either sign) to the firm. It is unrealistic to assume that the firm can be taxed by an unlimited amount or receive a very large bonus from the regulator. When there are limited liability restrictions, the first-best outcome cannot be implemented and the regulated firm can use this information differential in her advantage.

By the Revelation Principle, the regulator may resort, without loss of generality, to a self-selecting menu of contracts that induces truthful revelation of types. Such contracts must then satisfy the following incentive compatibility restrictions

$$
\begin{aligned}
U_{LL} &\geq U_{HL} + \Phi_1(e_{HL}), && \text{(1.17a)}\\
U_{HL} &\geq U_{LL} - \Phi_2(e_{LL}), && \text{(1.17b)}\\
U_{LH} &\geq U_{HH} + \Phi_1(e_{HH}), && \text{(1.17c)}\\
U_{HH} &\geq U_{LH} - \Phi_2(e_{LH}), && \text{(1.17d)}
\end{aligned}
$$

where $\Phi_1(e) = \Psi(e) - \Psi(e - 2\sigma_\theta)$ and $\Phi_2(e) = \Psi(e + 2\sigma_\theta) - \Psi(e)$ are strictly positive and increasing functions (see Appendix A.2 for a derivation of these

restrictions). By inducing revelation of θ and by observing the realized cost, the regulator can deduce with certainty, and therefore contract, the effort level. Proposition 1 below characterizes the optimal regulation in this setting, which is the maximization of (1.8) subject to (1.3), (1.6) and (1.17) above.

Proposition 1 (Second-Best (SB) Policy)**.** *Under full commitment and asymmetric information the optimal regulatory policy induces the first-best effort and output levels in all states except* (θ_H, m_H)*, in which*

$$\begin{aligned}\hat{e}_{HH} &= e^*_{HH} - 4\sigma_\theta \frac{\lambda}{1+\lambda}\frac{\psi}{\psi-(1+\lambda)}\frac{1-\rho}{1+\rho}\frac{\gamma}{\gamma+\sigma_m},\\ \hat{q}_{HH} &= q^*_{HH} - 4\sigma_\theta\lambda\frac{\psi}{\psi-(1+\lambda)}\frac{1-\rho}{1+\rho}\frac{\gamma}{\gamma+\sigma_m}\\ &= \psi\left[\hat{e}_{HH} + 4\sigma_\theta\frac{\lambda}{1+\lambda}\frac{1-\rho}{1+\rho}\frac{\gamma}{\gamma+\sigma_m}\right].\end{aligned}$$

The firm's rent is zero only in the (θ_H, m_H) *state, and its expected value is*

$$\mathbb{E}[\hat{U}] = I\frac{1+r_f}{1+\gamma/\sigma_m} + (\hat{q}_{HH} - \psi\sigma_\theta)(1-\rho)\frac{\gamma\,\sigma_\theta}{\gamma+\sigma_m}.$$

The systematic risk of the firm's cash flows are slightly higher under asymmetry, as some profit is received when $\hat{m} = m_H$*. Assuming* $I = q^*_{LH}$*, we have*

$$\beta_{SB} = \beta_{FB} + (1-\rho)\frac{\sigma_\theta}{\gamma+\sigma_m}\frac{\hat{e}_{HH}-\sigma_\theta}{e^*_{LH}}. \tag{1.18}$$

Proof. See Appendix A.2. ∎

When cost is privately observed by the firm, the regulator must leave some rent to the efficient firm to elicit that information even in good states of the economy, so $\hat{U}_{LH} > U^*_{LH} = 0$. As long as $\hat{e}_{HH} > \sigma_\theta$, the equilibrium profile of state-contingent profits raises the firm's systematic risk vis-a-vis the complete information case. This effect, however, is not large enough to result in a positive beta when we consider typical values for (1.18). See Section 1.4 for details.

Under incomplete information the regulator faces the usual trade-off between rent extraction and allocative efficiency at each realization of $\tilde{m}$. The efficient firm will always enjoy some rent, but the lower bound on the firm's value imposed by (1.6) creates the need to leave some rent to the inefficient firm as well, as informational rents may not be enough to accomplish this goal. The cheapest way to do this is to use U_{HL} to cover the financing gap, as doing otherwise would create systematic risk in the cash flows (so

$\hat{U}_{HH} = 0$ in equilibrium). In order to minimize the expected cost of rents, output and effort are distorted downwards in the high-cost / high-demand state, lowering the necessary U_{LH} to induce the efficient firm to reveal its cost truthfully. Also, notice that $\lim_{\rho \to 1} \hat{e}_{HH} = e^*_{HH}$, i.e., allocative distortion decreases with ρ. When cost is highly correlated with the aggregate shock, it is more likely that either (θ_L, m_L) or (θ_H, m_H) happens, so informational rent would naturally be counter-cyclical. And there is some $\bar{\rho} \in (-1, 0)$ such that if $\rho \leq \bar{\rho}$ production in the (θ_H, m_H) state would be shut down altogether. In that case, states (θ_H, m_L) and (θ_L, m_H) would be more likely and rents pro-cyclical, so $\partial\beta_{SB}/\partial\rho < 0$ in (1.18) above.

The implicit price of the good is similar to (1.11) in all states but the (θ_H, m_H), in which price is higher to limit the rent the efficient firm would extract under strong demand.

Under the policy described in Proposition 1, the highest effort level takes place in the (θ_L, m_H) state, like the complete information case, so condition (1.16) is still sufficient to guarantee a positive marginal cost in equilibrium $(e^*_{LH} < \theta_L)$. But the sufficient condition for $\hat{e}_{jk} > 0$ for all j, k is different, because $\hat{e}_{HH}$ may be severely reduced for incentive reasons and end up being lower than e^*_{HL}. A sufficient condition for $\min\{e^*_{HL}, \hat{e}_{HH}\} > 0$ is an upper bound for ψ,

$$\psi < \frac{[d(1 + \nu\sigma_m) - (1 + \lambda)\theta_H](1 + \lambda)(1 + \rho)(\gamma + \sigma_m)}{4\sigma_\theta(1 - \rho)\gamma\lambda}. \tag{1.19}$$

Then, we denote by

$$\hat{\mathcal{E}} = \{(d, \psi) \in \mathbb{R}^2_+ \mid (1.15), (1.16) \text{ and } (1.19) \text{ are valid}\}$$

the set of feasible economies under incomplete information. Conditions and it is straightforward to verify that $\hat{\mathcal{E}} \subset \mathcal{E}^*$.

Regulation Without Cost Observability (WC)

The previous section assumed that realized cost can be observed with certainty by all parties. This is a crucial assumption that allows the regulator to correctly infer (and contract) the level of effort employed by the firm by simply inducing revelation of the efficiency parameter θ. The observation of costs however may be an strong assumption in certain industries. Indeed, the pioneering work of Baron and Myerson (1982) is specifically based on this sort of informational advantage of the firm. Lack of cost observability restricts the regulator to a smaller set of instruments, which in our model would consist of only a transfer and output pair, or equivalently, the menu

$$\{(U_{jk}, q_{jk})_{j,k=L,H}\}. \tag{1.20}$$

In this context, individual rationality requires the transfer must be high enough to cover all costs incurred, so the *ex-post* rent become $U_{jk}(q,e) = t_{jk} - C(q,e,\theta_j) - \Psi(e)$ as costs cannot be reimbursed explicitly. As the contract specifies only a transfer and a quantity, the firm will privately choose the level of effort $E(q) \equiv \arg\max_e U_{jk}(q,e) = q/\psi$, and the truth-telling conditions, derived in Appendix A.2, are given by

$$U_{LL} \geq U_{HL} + 2\sigma_\theta\, q_{HL}, \tag{1.21a}$$
$$U_{LH} \geq U_{HH} + 2\sigma_\theta\, q_{HH}, \tag{1.21b}$$
$$U_{HL} \geq U_{LL} - 2\sigma_\theta\, q_{LL}, \tag{1.21c}$$
$$U_{HH} \geq U_{LH} - 2\sigma_\theta\, q_{LH}. \tag{1.21d}$$

Ex-post welfare is re-written as $W_{jk}(q,U) \equiv W_{jk}(q,E(q),U)$ and the optimal regulation is a set of contracts (1.20) that maximizes the expected welfare subject to (1.3), (1.6) and (1.21), and is summarized in Proposition 2.

Proposition 2. *Under full commitment, asymmetric information and cost unobservability, output is distorted in the (θ_H, m_H) state to a level similar as described in Proposition 1, so*

$$\underline{q}_{jk} = \hat{q}_{jk},$$

for all $j,k = L,H$. However, since the firm will always exert the first-best level of effort, $\underline{e}_{jk} = \underline{q}_{jk}/\psi$ for all $j,k = L,H$. In particular, $\underline{e}_{HH} = \underline{q}_{HH}/\psi > \hat{e}_{HH}$. The firm's rent is null only in the (θ_H, m_H) state and has an expected value of

$$\mathbb{E}[\underline{U}] = I\frac{1+r_f}{1+\gamma/\sigma_m} + \underline{q}_{HH}(1-\rho)\sigma_\theta\frac{\gamma}{\gamma+\sigma_m},$$

which is higher than $\mathbb{E}[\hat{U}]$. The systematic risk suffers an additional increase when cost is not observable:

$$\begin{aligned}\beta_{WC} &= \beta_{FB} + (1-\rho)\frac{\sigma_\theta}{\gamma+\sigma_m}\frac{\hat{q}_{HH}}{q^*_{LH}} > \beta_{FB},\\ &= \beta_{SB} + (1-\rho)\frac{\sigma_\theta}{\gamma+\sigma_m}\frac{\hat{q}_{HH} - \psi(\hat{e}_{HH}-\sigma_\theta)}{q^*_{LH}} > \beta_{SB}.\end{aligned}$$

Proof. See Appendix A.2.1. ■

When costs are observable, truthful revelation of θ allows the regulator to contract the firm's privately observed effort. As shown in Proposition 1, this flexibility allows the regulator to implement a quantity $\hat{q}_{HH} > \psi\,\hat{e}_{HH}$ that is less distorted, closer to the first-best level that would prevail in that

state. The regulator can, to some extent, decouple the distortion of effort for rent extraction purposes from the distortion of output. But when costs are not observable, the firm will always choose to apply the efficient level of effort $\underline{e}_{jk} = \underline{q}_{jk}/\psi$, so rent extraction is more costly in terms of welfare as it has to be carried out directly on output.

In terms of restrictions on exogenous parameter for a meaningful equilibrium, condition (1.16) is still sufficient for $\underline{e}_{jk} < \theta_j$ at all j, k. For a positive effort in every state, as $\underline{e}_{HH} > \hat{e}_{HH}$, the upper bound for ψ is *less* strict than (1.19), so these two set of restrictions are sufficient for a well-behaved solution when cost is not observed.

Price-Cap Regulation (PC)

In addition to not being able to observe realized cost and its constituents, the regulator is usually forbidden to operate direct transfers with the firm. As a matter of fact, this is the rule rather than the exception in regulatory (as opposed to procurement) environments. In this case the regulator may resort to a price-cap system in which the firm is allowed to charge a state-independent maximum price p per unit directly from customers. In this system there is no revelation mechanism, and given a price-cap p the firm will enjoy a cash flow of

$$U_{jk}(p, e) = (p - \theta_j + e)q_k(p) - \psi e^2/2, \tag{1.22}$$

where $q_k(p) \equiv q(p, m_k)$ is the demand function. At the prevailing price p the firm will optimally choose to exert effort $E_k(p) = \arg\max_e U_{jk}(p, e) = q_k(p)/\psi$, and we denote $U_{jk}(p) \equiv U_{jk}(p, E_k(p))$ the cash flow under the price-cap p and efficient effort. Individual rationality is then

$$U_{jk}(p) \geq 0 \quad \text{for all } j, k \tag{1.23}$$

and the financing condition becomes

$$V(p) \equiv \sum_j \sum_k P(\theta_j, m_k)U_{jk}(p)(1 - \gamma/m_k) \geq I. \tag{1.24}$$

Under this regime, the regulator does not reimburse the firm the productions costs, so an utilitarian social welfare is more appropriately defined as

$$\mathcal{W}_{jk}(p) \equiv \int_p^{d(1+m_k\nu)} q_k(z)dz + \alpha U_{jk}(p),$$

where $\alpha \in [0, 1]$ is a preference parameter that reflects distributional concerns, i.e., the relative weight of profits vis-à-vis consumer surplus in the

social welfare. As noted by Armstrong and Sappington (2007), a regulatory environment with an objective such as $\mathcal{W}_{jk}(p)$ and $\alpha < 1$ results in policies that are similar in qualitative terms to those obtained in a set-up where welfare is (1.1) and $\lambda > 0$. Also, in (1.1) profit has a weight equal to $-\lambda$, so the planner strictly dislikes leaving rent to the firm. On the other hand, any $\alpha > 0$ in $\mathcal{W}_{jk}(p)$ would attach positive weight to profits. Hence, in order to retain the comparability of the different regimes considered, and to avoid the inclusion of one additional parameter with negligible qualitative effects on policy, we set $\alpha = 0$ to obtain

$$\mathcal{W}_k(p) \equiv q_k(p)^2. \tag{1.25}$$

The optimal price-cap policy is a pair $\{p, I\}$ that maximizes the expected value of (1.25) subject to (1.23), (1.24) and the capacity constraints

$$q_k(p) \leq I \quad \text{for all } k. \tag{1.26}$$

From (1.26) we immediately deduce that $q_L(p) \leq I$ is redundant and that $q_H(p) = I$ at any optimum. Also, in the appendix we show that if $\theta_H < p < d(1 + \sigma_m \nu)$, then $U_{HL}(p) \leq U_{jk}(p)$ for all j and k, so we can replace the four IR constraints in (1.23) by $U_{HL}(p) \geq 0$. We verify ex-post whether the optimal price is indeed in the range $(\theta_H, q_H^{-1}(0))$ and summarize the optimal price-cap regime in Proposition 3 below.

Proposition 3. *The optimal price-cap regime is the pair $\{p^{pc}, I^{pc}\}$ such that p^{pc} maximizes*

$$\mathbb{E}\mathcal{W}(p) \equiv \sum_j \sum_k P(\theta_j, m_k)\mathcal{W}_k(p)$$

subject to $V(p) \geq q_H(p)$ and $U_{HL}(p) \geq 0$. Also,

$$I^{pc} = q_H(p^{pc}).$$

The price-cap is feasible when effort and marginal cost are positive in every state, which requires $0 < E_L(p^{pc})$ and $E_H(p^{pc}) < \theta_L$ or

$$d(1 + \sigma_m \nu) - \psi\, \theta_L < p^{pc} < d(1 - \sigma_m \nu). \tag{1.27}$$

Proof. See Appendix A.2.2. ∎

In the simulations at Section 1.4, the optimal price-cap is always the lowest positive price that solves $V(p) = q_H(p)$, and the IR and feasibility bounds are always satisfied at this price. While in this case a closed-form solution for p^{pc} is available, it is too intricate to be given explicitly here.

As a consequence, the full expression for β_{PC}, the systematic risk of a firm under the price-cap, is also too involved and will not be given here. It is also very difficult to obtain the exact regions for the parameters (d, ψ) such that the price-cap solution is feasible. Given these difficulties, we will simulate optimal policies in economies that belong to $\hat{\mathcal{E}}$ and will verify ex-post whether the resulting policies generate sensible equilibrium values, with positive effort and marginal cost at any state.

Profit-Sharing Regulation (PS)

A price-cap regime, while providing strong incentives for cost reduction, generally allows the firm to earn substantial profits. This phenomena has been used to justify profit-sharing systems, where the company is given a price-cap p but a fraction $\tau \in [0, 1]$ of earnings are taxed by the regulator. Compared with the price-cap regime, a profit-sharing rule gives less incentive for cost reduction (i.e., the power of the regulatory contract is lower) but the ability to extract rent from the firm is higher. It also requires the observability of profits and the possibility of transfers from the firm to the regulator. Under a profit-sharing rule, the firm's net profit is

$$U_{jk}(p, e, \tau) = (1 - \tau)(p - \theta_j + e)q_k(p) - \psi e^2/2, \tag{1.28}$$

as the disutility of effort is not observed by the regulator and thus cannot be part of the taxable income. When faced with a profit-sharing policy $\{p, \tau, I\}$, the firm chooses to exert $E_k(p, \tau) \equiv \arg\max_e U_{jk}(p, e, \tau) = (1 - \tau)q_k(p)/\psi$, and we denote by $U_{jk}(p, \tau) \equiv U_{jk}(p, E_k(p, \tau), \tau)$ the net profit under the optimal effort. Notice that the optimal effort decreases with the tax rate. The limited liability restrictions become

$$U_{jk}(p, I) \geq 0 \quad \text{for all } j, k \tag{1.29}$$

As in the price-cap case, one can show that $U_{HL}(p, \tau) \leq U_{jk}(p, \tau)$ for all j, k and (p, τ). Therefore, the restrictions in (1.29) can be replaced by $U_{HL}(p, \tau) \geq 0$. The financing constraint of the firm under the profit-sharing regime is given by

$$V(p, \tau) \equiv \sum_j \sum_k P(\theta_j, m_k) U_{jk}(p, \tau)(1 - \gamma/m_k) \geq I. \tag{1.30}$$

The total amount of tax revenue raised is $T_{jk}(p, \tau) = \tau(p - \theta_j + e)q_k(p)$, which when used to decrease taxes elsewhere in the economy, adds $\lambda T_{jk}(p, \tau)$ to the net welfare of consumers. The ex-post welfare under a profit-sharing rule is then given by

$$\mathcal{W}_{jk}(p, \tau) \equiv q_k(p)^2 + \lambda T_{jk}(p, \tau). \tag{1.31}$$

The optimal profit-sharing policy is the set $\{p^{ps}, \tau^{ps}, I^{ps}\}$ that maximizes the expected value of (1.31) subject to $U_{HL}(p,\tau) \geq 0$ and $V(p,\tau) \geq I$ and the capacity restraints (1.26). Note however that we have $q_L(p) < q_H(p) = I$ at the solution, so we can search for the optimal price and tax-rate only. Proposition 4 below characterizes the optimal profit-sharing regime.

Proposition 4. *The optimal profit-sharing regime is the triple* $\{p^{ps}, \tau^{ps}, I^{ps}\}$ *such that* p^{ps} *and* τ^{ps} *maximize*

$$\mathbb{E}\mathcal{W}(p,\tau) \equiv \sum_j \sum_k P(\theta_j, m_k)\mathcal{W}_{jk}(p,\tau)$$

subject to $V(p,\tau) \geq q_H(p)$, $U_{HL}(p,\tau) \geq 0$ *and* $0 \leq \tau \leq 1$. *Also,*

$$I^{ps} = q_H(p^{ps}).$$

The price-cap is feasible when effort and marginal cost are positive in every state, which requires $0 < E_L(p^{ps}, \tau^{ps})$ *and* $E_H(p^{pc}, \tau^{ps}) < \theta_L$ *or*

$$d(1+\sigma_m\nu) - \frac{\psi\,\theta_L}{1-\tau^{ps}} < p^{ps} < d(1-\sigma_m\nu). \tag{1.32}$$

Proof. Ommited. ■

Although it can be easily obtained numerically, the profit-sharing policy is intractable to be given explicitly. The restrictions on the exogenous parameters that guarantee a well-behaved solution are also difficult to obtain in closed form. Hence, we follow the same approach used in the price-cap system and will simulate policies in economies that belong to the more restricted set $\hat{\mathcal{E}}$, and will very ex-post whether equilibrium values are feasible and meaningful.

With the profit-sharing regime described above we can extend our analysis of regulated betas to a very important dimension of the regulatory practice: the power of the regulatory contract and its impact on the risk of the firm. In moral hazard situations, power is related to the agent's willingness to exert the hidden action under the terms of the contract. In the regulatory context it is the incentive for cost reduction implied in the price rule, and is usually measured by the percentage of the savings that the shareholders can keep for themselves in the form of profits. In a price-cap regime, power is maximal (i.e., equals 1) as the company cashes in every dollar of cost savings. At the other extreme is the cost-plus (rate of return) regulation, in which savings are

transferred to customers in the form of a lower price in the next rate revision[3], so power is minimal (equals 0). The power in a profit-sharing contract stands between these two extremes, because the company keeps $100(1-\tau)\%$ of cost reductions from a marginal increase in effort, so power is exactly the fraction $(1-\tau)$. It is interesting to see how the systematic risk of a company under profit-sharing varies with τ. In fact, as τ is endogenous, in Section 1.4 we will investigate how the optimal tax rate changes when exogenous quantities like cost and demand uncertainty σ_θ and ν vary, and the associated response from β_{PS}, the systematic risk of the firm under profit-sharing regulation.

1.4 Numerical Analysis

Some of the regulatory models discussed do not yield closed-form solutions for the equilibrium policy variables and beta, so it is difficult to compare their different quantitative properties. It is necessary to calibrate the relevant parameters so we can have a better understanding of how the fundamental characteristics of the industry such as cost and demand uncertainty, σ_θ and ν, affect the systematic risk of the regulated firm. With the numerical simulations we can also show how beta responds to increasing limitations on the regulatory environment, as we gradually move from a first-best policy in which transfers are allowed to a second-best contract without transfers. We can also shed some light on the relationship between beta and contract power, using the profit-sharing regime described in Section 1.3.2 as the benchmark model.

Some parameters in our model are related to aggregate and macroeconomic variables, so are exogenous to the regulated industry and can be directly evaluated by plugging in historical averages or estimates from previous studies. For the risk-free rate r_f, the market risk premium γ and the volatility of the market portfolio σ_m, we use historical averages observed in the U.S. capital markets. In Table A.2 we show the returns, averaged for 5-year holding periods, of intermediate-term Treasury bonds and a portfolio of large stocks, proxies for the risk-free rate and market portfolio, respectively.

[3] Under cost-plus regulation the firm can keep the savings obtained *within* rate revisions, so there is some incentive for efficient production as long as rates are kept unchanged for some period. However, substantial cost savings are likely to trigger a revision process, effectively diminishing this incentive. The price-cap, on the other hand, is valid for several years (usually five) and the timing of the revision process is exogenous and set *ex-ante*. In practice however, this distinction is blurred, as the rate revision process under the cost-plus may be managed (delayed or anticipated) by the political lobbies involved, and as it is also hard for the firm under a price-cap to politically sustain abnormal, and yet perfectly legal, profits during a long period of time.

The historical 5-year averages in Table A.2 are compatible with commonly reported annual returns of 6% for the risk-free rate and 12% for the stock market with a 20% standard deviation. We set $r_f = 30\%$, $\gamma = 35\%$ and $\sigma_m = 50\%$ in the simulations.

The shadow cost of public funds λ is set at a base value of 0.3, considered a reasonable estimate for a developed country (see Ballard, Shoven and Whalley, 1985; Hausman and Poterba, 1987) and used by Gasmi, Kennet, Laffont and Sharkey, 2002 in their analysis of the telecommunications sector in the United States.

With the macro parameters defined as above, we now discuss the industry-specific ones. First, we normalize the expected value of the base marginal cost at $\mu_\theta = 1$ so it will serve as the numeraire for the other industry parameters, which will then be interpreted as multiples or percentages of the basic marginal cost. We work with the intervals $\sigma_\theta \in (0, 0.20]$ and $\nu \in (0, 0.30]$, which probably cover a fair portion of the actual degree of cross-sectional variation of cost and demand uncertainty that might exist in a regulated sector. When $\sigma_\theta = 0.20$, there is a difference of approximately 40% in the marginal cost between the least and most efficient firms. And since $\sigma_m = 50\%$, when $\nu = 0.30$ the actual demand vary between $\pm 15\%$ around the expected value, which is also a large level of uncertainty in regulated sectors of essential services, where demand tends to be relatively stable over time.

We consider $\rho \in [-0.50, 0.50]$ a representative interval, as the firm's efficiency does not seem to depend too much on aggregate shocks in the real world. Although economies of scale and scope and input costs are partially influenced by the business cycle, there are many factors that have a deep impact on efficiency that are not related to the performance of the economy (managerial talent and experience, access to new technologies, etc.). We are not aware of any empirical study that estimated such a parameter for a regulated industry, so the chosen interval is arbitrary. But in our model, in spite the fact that a $\rho \neq 0$ does influence policy by changing the likelihood of higher costs and the cyclicality of informational rents, this is not an essential requirement for the qualitative results we obtained. There are still departures from the first-best allocation even when $\rho = 0$ under incomplete information. In conclusion, extreme values of ρ seem unlikely in practice, and we are able to retain most qualitative features of the different polices by working with this restricted interval.

With σ_θ, ν, ρ and the macro parameters defined as above, we can characterize the set (d, ψ) of feasible economies $\hat{\mathcal{E}}$. Let $\psi_1(d, \sigma_\theta, \nu)$ denote the right-hand side of (1.16), so it is the lower bound for ψ in $\hat{\mathcal{E}}$ and an increasing function of all its arguments. Also, let $\psi_2(d, \sigma_\theta, \nu, \rho)$ denote the right-hand side of (1.19), so it is the upper bound for ψ in $\hat{\mathcal{E}}$ and increases with all

arguments except σ_θ. For each choice of d, the smallest possible range size of feasible values for ψ is $R_\psi(d) \equiv \psi_2(d, 0.20, 0, -0.50) - \psi_1(d, 0.20, 0.30)$, which strictly increases with d and is non-empty when $d > 2.33$. Then, we let

$$\begin{aligned} \psi_L(d) &\equiv \psi_1(d, 0.20, 0.30) + 0.1R_\psi(d), \\ \psi_H(d) &\equiv \psi_2(d, 0.20, 0, -0.50) - 0.1R_\psi(d), \\ \tilde{\psi}(d) &\equiv \psi_1(d, 0.20, 0.30) + R_\psi(d)/2, \end{aligned}$$

so that $\psi_L(d)$ is slightly above the highest lower bound, $\psi_H(d)$ is sligthly below the lowest higher bound, $\psi_H(d) - \psi_L(d) > 0$ is the largest possible feasible interval for ψ with $\tilde{\psi}(d)$ as its midpoint. These three levels of ψ are always within the bounds (1.16) and (1.19) for any values of the σ_θ, ν and ρ parameters, so they form a consistent basis for comparison that can be held fixed when simulating different regulatory systems, while providing a reasonable variation across the spectrum of possible values for the cost of inducing effort. Thus, we consider the set $\{\psi_L(d), \tilde{\psi}(d), \psi_H(d)\}$ for ψ in the simulations.

Finally, the demand intercept d must obey the lower bound (1.15), which is an increasing function of σ_θ and ν. Evaluating (1.15) at $\sigma_\theta = 0.20$, $\nu = 0.30$, $\lambda = 0.3$ and $\sigma_m = 0.50$, we conclude that any $d > 1.84$ is in $\hat{\mathcal{E}}$. We consider three levels for the demand intercept, based on the following rationale. Consider the optimal price-cap policy of Section 1.3.2, evaluated at the midpoints $\sigma_\theta = 0.10$, $\nu = 0.15$, $\psi = \tilde{\psi}(d)$ and with profits mildly pro-cyclical, $\rho = -0.25$. Under these assumptions, there is an optimal price-cap for each value of of the intercept, $p^{pc}(d)$. Figure 1.1 below plots the expected price-elasticity of demand evaluated at the optimal price-cap, $\mathbb{E}[\eta] = p^{pc}(d)/(d - p^{pc}(d))$, against d.

In Figure 1.1 above we see that typical price-elasticities found in regulated industries, e.g., 0.50, 0.25 and 0.05 in absolute value, are associated, in our model, with demand intercepts at approximately 7.38, 12 and 50, respectively. A range between -0.05 to -0.50 is a reasonable representation of actual price-elasticities of demand for essential goods. See, for instance, Taylor (1975), Taylor (1994) and Bohi and Zimmerman (1984) for empirical studies of demand in the telecommunications and energy sectors. Schmalensee (1989) and Gasmi, Ivaldi and Laffont (1994) set d as the numeraire and adjust θ to a chosen elasticity at $p = \theta$. We therefore consider $d \in \{7.38, 12, 50\}$ for the simulations.

We defined a wide range of values for the model parameters that are valid across all different regulatory regimes reviewed in Section 1.3, and we are now able to evaluate how β responds to different industry characteristics

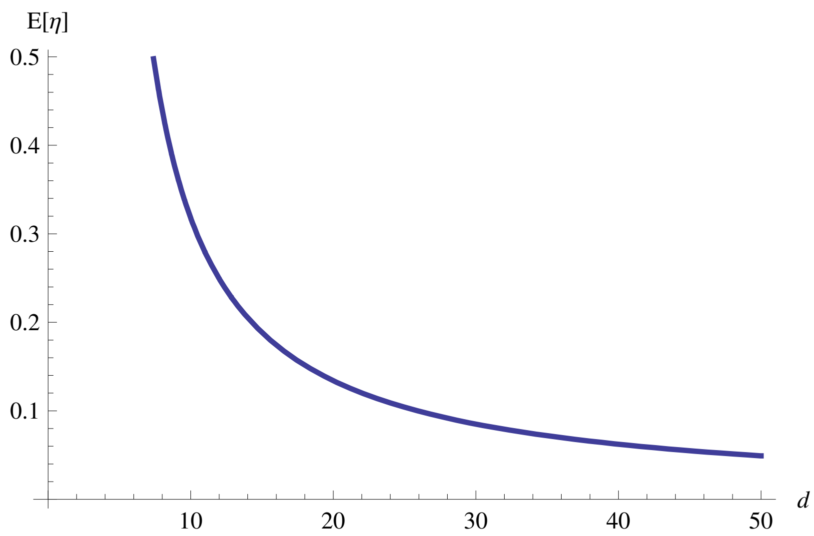


Figure 1.1: Expected price-elasticity of demand under the price-cap.

and regimes on a consistent basis. We obtained analytical results for just a subset of possible regimes, so in order to have a better understanding of the properties of each one we simulate[4] the optimal policy and calculate the equilibrium β at different levels of the parameters. In Section 1.4.1 we gauge the sensitivity of β to industry parameters, i.e., related to demand (d and ν) or technology (ψ, σ_θ and ρ). Next, in Section 1.4.2 we fix d and ψ at average levels and compare how the different regulatory regimes affect β when market uncertainty parameters (ν, σ_θ and ρ) change. Finally, in Section 1.4.3 we focus on how the (endogenous) contract power and beta vary with market uncertainty parameters in a profit-sharing regime. In every scenario, we verified *ex-post* that the conditions (1.27) and (1.32) are satisfied, so the equilibrium price-cap and profit-sharing regimes are feasible throughout all simulations. Table A.3 summarizes the restrictions imposed on the endogenous parameters by the model (in the first column), by the equilibrium conditions in the SB regime (second column) and by the additional refinements discussed previously (in the third and last column).

1.4.1 Sensitivity of β to Industry Parameters

In the complete information case, the firm has a systematic risk that is invariant to any policy variable and depends only on exogenous macro parameters. But in more restricted contracting environments, β is affected by cost and

[4]All numerical exercises were conducted with the Mathematica software.

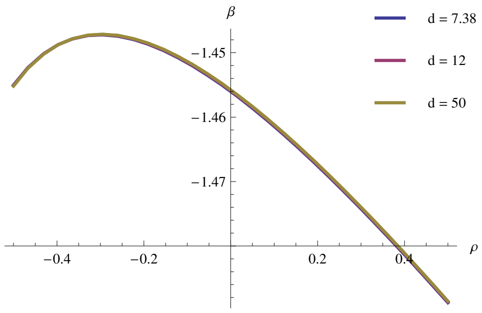
demand conditions and the degree of uncertainty in those parameters.

We simulated the value and responsiveness of β to changes in industry conditions in the following manner. Taking as given the values of the macro parameters discussed previously and for each regulatory system described in Section 1.3, we performed two sets of scenarios. In the first, we held fixed the demand parameter at its mid-point ($d = 12$) and calculated the equilibrium β against each uncertainty parameter ν, σ_θ and ρ and level of the cost of effort, $\psi_L(12)$, $\tilde{\psi}(12)$ and $\psi_H(12)$, while keeping the remaining parameters fixed at their average (or representative) values: $\nu = 0.15$, $\sigma_\theta = 0.10$ and $\rho = -0.25$. In the second set we simulated β against the uncertainty parameters in the same manner, but across all levels of $d \in \{7.38, 12, 50\}$ while keeping the disutility of effort constant at its average value, $\psi = \tilde{\psi}(d)$. In Figures 1.2 to 1.5 we plot the results for each regulatory regime and simulation scenario.

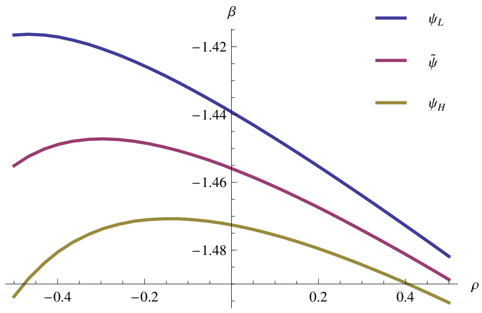
In the second-best regulation (SB), shown in Figure 1.2, the firm's beta is slightly larger then β_{FB} across most of the parameter space, but it is fairly insensitive to industry characteristics. When the desutility of effort is held at its average value (panels 1.2a, 1.2c and 1.2e), wide variation in price-elasticity (different values of d) has a negligible impact on β, and changes in cost uncertainty (σ_θ) or market correlation (ρ) do not cause a change of more than 6 decimal points in β. Panels 1.2(b), 1.2(d) and 1.2(f) show that the impact of ψ is slightly more pronounced in the SB regime, but the influence of the remaining factor are still very small. In the second-best regime, in spite of having an informational disadvantage against the firm, the regulator is capable of allocating rents in a way that almost completely mitigates aggregate risk.

The situation does not change much when we remove the regulator's ability to observe *ex-post* costs. Under the WC regime, β is almost completely insensitive to demand elasticity and to the cost of inducing effort, as Figure 1.3 shows. As it might be expected, cost-related parameters have the larger (but still small in absolute magnitude) effect on β: raising cost uncertainty (standard deviation) from 1% to 20% raises β in 25 decimal points (regardless the choice of d or ψ); and as cost efficiency becomes more pro-cyclical (decreasing ρ from 0.5 to -0.5), β raises by a full decimal point. The end result however still resembles less restricted regimes such as the first and second-best, as β_{WC} is never higher than -1 in any of the simulated scenarios. For the SB and WC systems, there does not exist a combination of parameters that significantly departures their equilibrium betas from β_{FB} (of approximately -1.50 given our calibrated macro parameters).

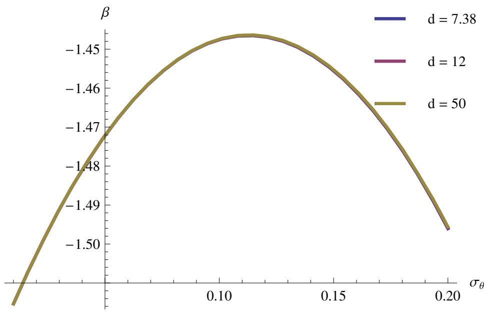
However, when the ability of the regulator to operate direct transfers is re-



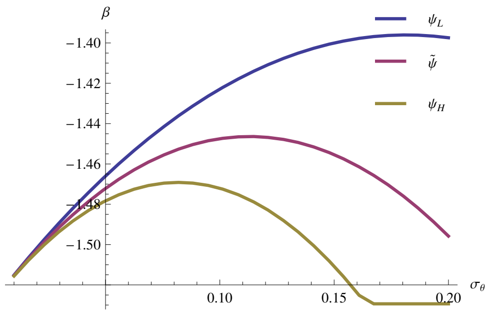
(a) $\psi = \tilde{\psi}(12)$, $\nu = 0.15$ and $\sigma_\theta = 0.10$



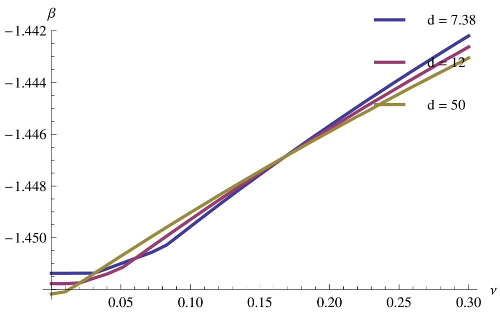
(b) $d = 12$, $\nu = 0.15$ and $\sigma_\theta = 0.10$



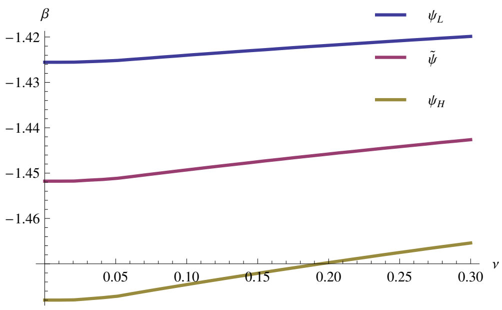
(c) $\psi = \tilde{\psi}(12)$, $\nu = 0.15$ and $\rho = -0.25$



(d) $d = 12$, $\nu = 0.15$ and $\rho = -0.25$



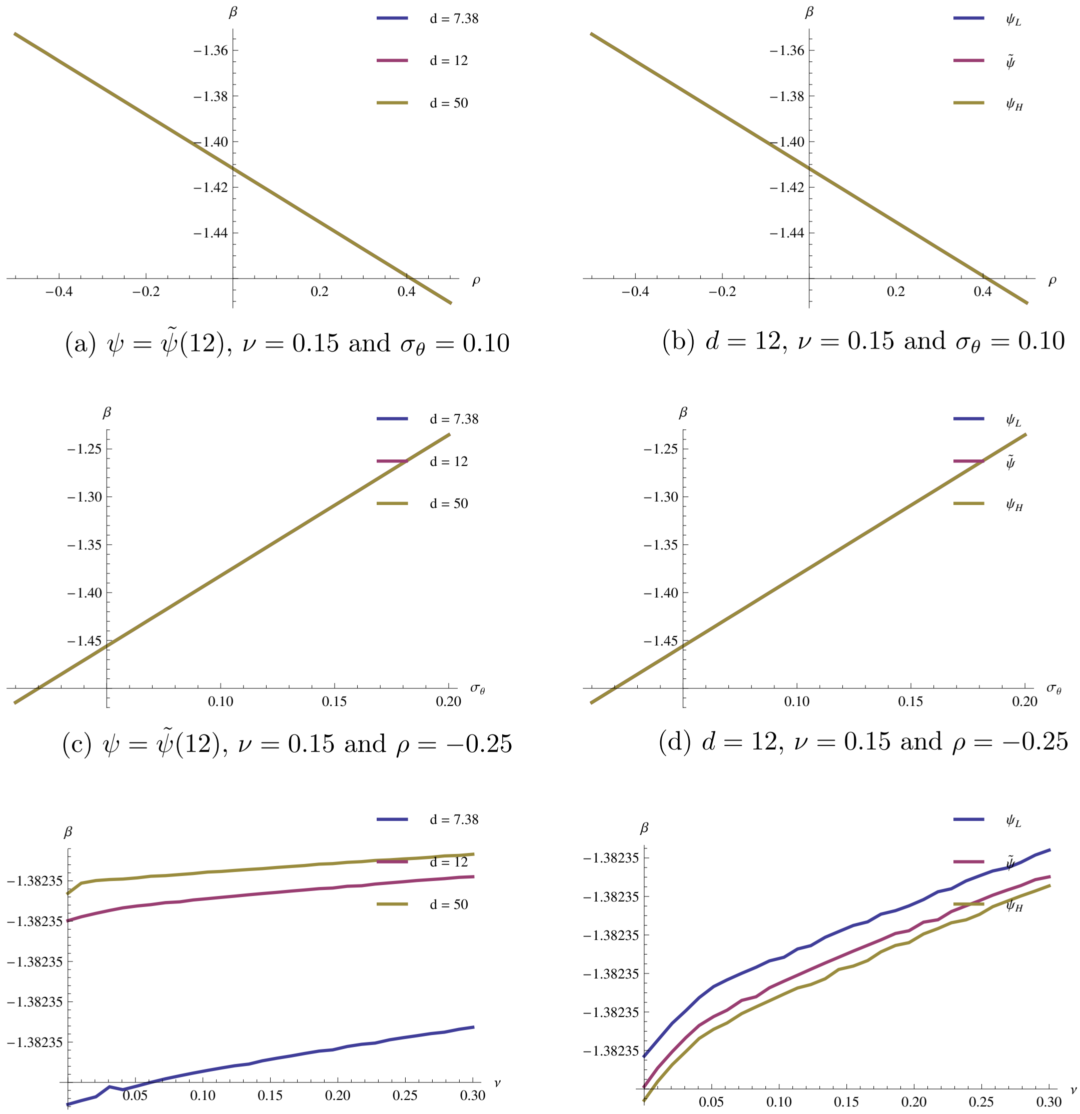
(e) $\psi = \tilde{\psi}(12)$, $\sigma_\theta = 0.10$ and $\rho = -0.25$



(f) $d = 12$, $\sigma_\theta = 0.10$ and $\rho = -0.25$

Figure 1.2: Sensibility of β to industry parameters in the SB regime.

duced, the risk profile of firm changes substantially. In Figures 1.4 and 1.5 we show the results for the profit-sharing (PS) and price-cap (PC) regimes. First and foremost, β is positive across the entire parameter space in both regimes. Second, β_{PS} and β_{PS} monotonically increase with cost and demand uncertainty, when profits become more pro-cyclical and with the price-elasticity of



(a) $\psi = \tilde{\psi}(12)$, $\nu = 0.15$ and $\sigma_\theta = 0.10$

(b) $d = 12$, $\nu = 0.15$ and $\sigma_\theta = 0.10$

(c) $\psi = \tilde{\psi}(12)$, $\nu = 0.15$ and $\rho = -0.25$

(d) $d = 12$, $\nu = 0.15$ and $\rho = -0.25$

(e) $\psi = \tilde{\psi}(12)$, $\sigma_\theta = 0.10$ and $\rho = -0.25$

(f) $d = 12$, $\sigma_\theta = 0.10$ and $\rho = -0.25$

Figure 1.3: Sensibility of β to industry parameters in the WC regime.

demand[5], as we would expect. Third, the simulated betas are very close to the observed ones reported in Table A.1, specially when we consider average

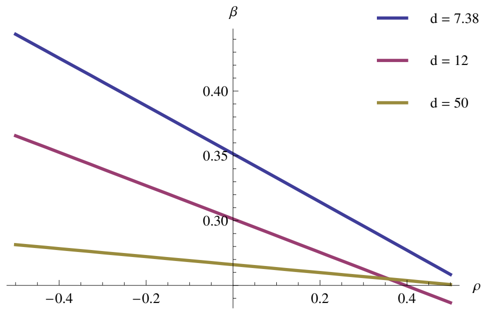
[5]For all practical purposes, we can safely ignore the regions with $\rho > 0$ in Figures 1.4(a) and 1.4(b) in which β_{PS} is larger for $d = 50$ than $d = 12$, i.e., when the expected price-elasticity of demand is lower. This result reflects a rare situation, in which the profits of a regulated enterprise are strongly counter-cyclical.

levels of demand and cost uncertainty in the simulation. Lastly, the numerical value of cost of inducing effort does not have a significant influence on β_{PS} or β_{PC}. This is a fortunate result, because although ψ plays a very important *qualitative* role in the regulatory game, it is by far the most abstract parameter in the model and has no natural counterpart in the real world from where we can search for a representative numerical value. For the simulation exercise, we identified upper and lower bounds for ψ that are consistent with meaningful equilibrium values for all regulatory regimes simultaneously, and found that β does not change significantly as we raise ψ from ψ_L to ψ_H.

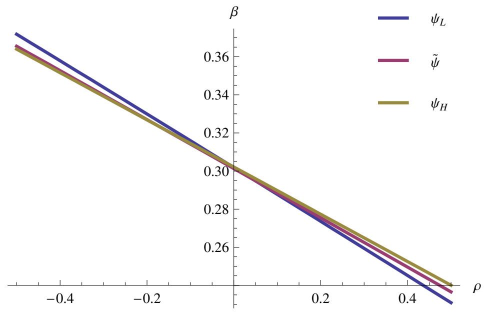
If we consider the different forms of second-best regulation (SB and WC) as one group, and the profit-sharing and price-cap systems as another, and compare their major numerical properties, two clear distinctions emerge. First, the responsiveness of β to changes in the market fundamentals is significantly higher under the regimes where transfers are restricted. For example, the largest variation in β we were able to generate in the first group was 25 decimal points, as can be seen in Figures 1.3(c) and 1.3(d). On the other hand, there is a difference of approximately 80 decimal points among the lowest and highest β in some scenarios of the second group, as shown in Figures 1.4(e) and 1.5(e). Second, while in the first group cost-related variables (σ_θ and ρ) were the most influential on β, in the second group demand variables (ν and d) play the major role. It seems that direct transfers are an effective way to mitigate demand risk, and in the absence of such an instrument the systematic risk of the regulated entity is mainly influenced by demand uncertainty and price-elasticity.

1.4.2 Sensitivity of β to the Contract Space

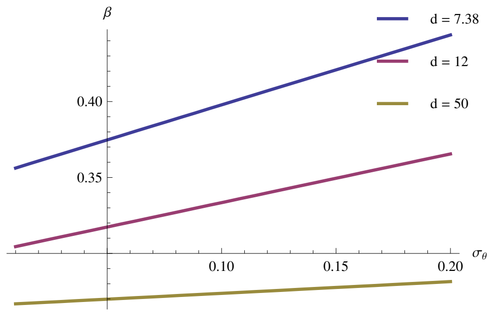
The regulatory systems reviewed in Section 1.3 can be sorted in increasing order of restrictiveness in the contracting environment. The first-best (FB) policy is the less restricted, as the regulator has complete information about the firm's operating environment, can contract on all relevant variables (q, e and U) and use direct transfers with the firm. In the second-best (SB), the authority no longer observes the different components of marginal cost, but still has access to transfers and designs a mechanism that allows contracting on output, effort and cash-flow just like in the first-best. In the WC policy the regulator uses direct transfers but loses the ability to observe realized costs, so the contract must be written in terms of (q, U) only, as the firm chooses her effort privately. In the profit-sharing (PS) system, only one-way transfers (taxation of the firm) are allowed, and the instruments available



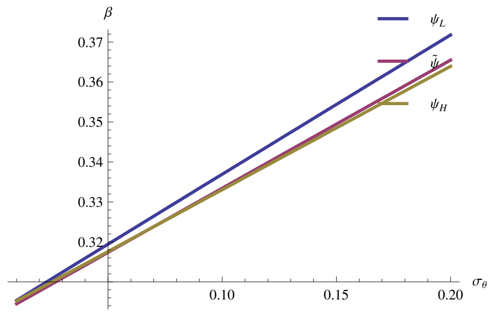
(a) $\psi = \tilde{\psi}(12)$, $\nu = 0.15$ and $\sigma_\theta = 0.10$



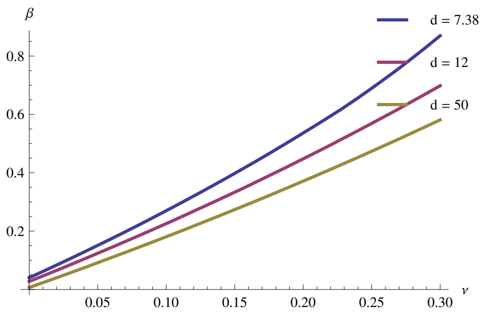
(b) $d = 12$, $\nu = 0.15$ and $\sigma_\theta = 0.10$



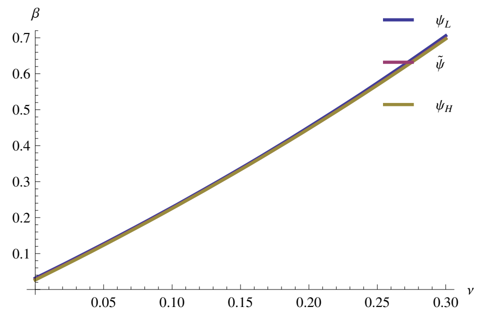
(c) $\psi = \tilde{\psi}(12)$, $\nu = 0.15$ and $\rho = -0.25$



(d) $d = 12$, $\nu = 0.15$ and $\rho = -0.25$



(e) $\psi = \tilde{\psi}(12)$, $\sigma_\theta = 0.10$ and $\rho = -0.25$

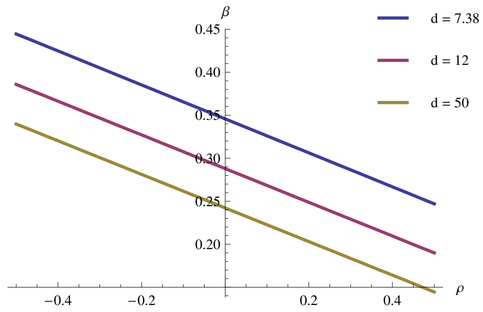


(f) $d = 12$, $\sigma_\theta = 0.10$ and $\rho = -0.25$

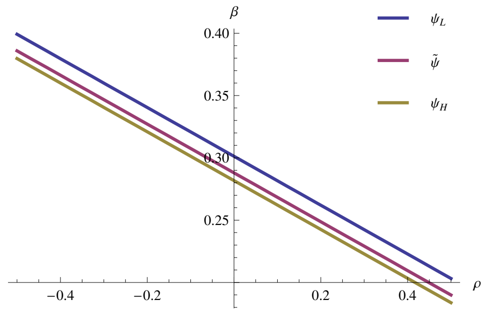
Figure 1.4: Sensibility of β to industry parameters in the PS regime.

are only a price-tax rate pair (p, τ) that is constant across states. Finally, the most restricted contract is the price-cap (PC), in which no transfers are allowed and the contract specifies a single ceiling price p to be charged by the firm in every state.

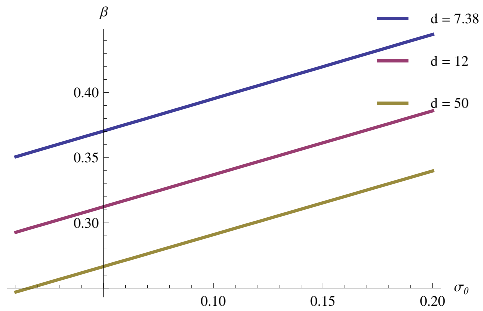
Given such a gradual composition of regulatory systems, it is natural to compare betas across these different systems and to attribute the possible



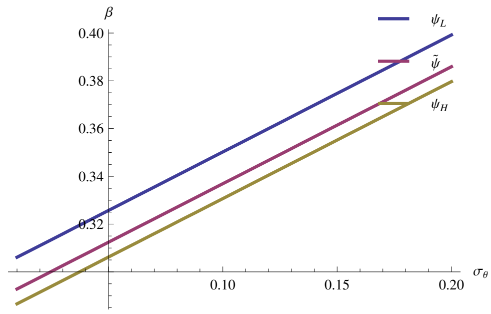
(a) $\psi = \tilde{\psi}(12)$, $\nu = 0.15$ and $\sigma_\theta = 0.10$



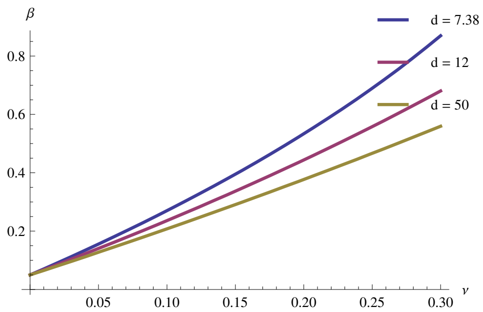
(b) $d = 12$, $\nu = 0.15$ and $\sigma_\theta = 0.10$



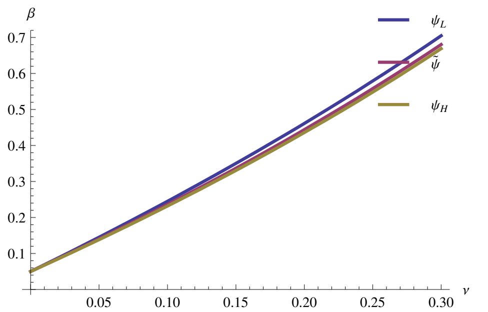
(c) $\psi = \tilde{\psi}(12)$, $\nu = 0.15$ and $\rho = -0.25$



(d) $d = 12$, $\nu = 0.15$ and $\rho = -0.25$



(e) $\psi = \tilde{\psi}(12)$, $\sigma_\theta = 0.10$ and $\rho = -0.25$



(f) $d = 12$, $\sigma_\theta = 0.10$ and $\rho = -0.25$

Figure 1.5: Sensibility of β to industry parameters in the PC regime.

discrepancies to the varying level of flexibility in the space of contracts. In order to do that, we fixed all industry parameters at their average, representative values, and calculated the equilibrium β under each policy and across the entire feasible range of each parameter at a time.

In each panel of Figure 1.6 we plot β_{FB}, β_{SB}, β_{WC}, β_{PS} and β_{PC} against one industry parameter while keeping the others fixed at their average levels.

Figure 1.6 makes it clear why it is meaningful to classify the different regimes in two classes, one in which transfers are unrestricted (FB, SB and WC) and the other with PS and PC, as we did at the end of Section 1.4.1. Policies that belong to the second group are the only ones capable of generating positive betas, in line with the results obtained in Section 1.4.1. We can also the previous finding that, at least when measuring at average parameter values, β varies the most in the second group and by changing the level of demand uncertainty.

Another general conclusion that can be drawn is that the risk profile of each regime can be sorted in the same ordering that we have classified the regimes themselves, i.e.

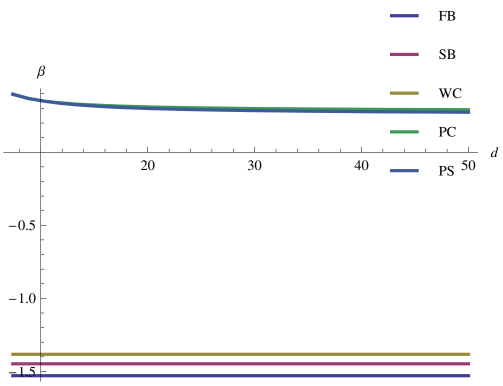
$$\beta_{FB} < \beta_{SB} < \beta_{WC} < 0 < \beta_{PS} \approx \beta_{PC}$$

so the simulation exercise suggests that contract flexibility, specially the ability to operate transfers, significantly affect the regulated business' β. However, once we endogenize the optimal sharing rule τ, we cannot tell whether the profit-sharing or the price-cap induce more risk to the business, as they appear to generate approximately the same profile of betas.

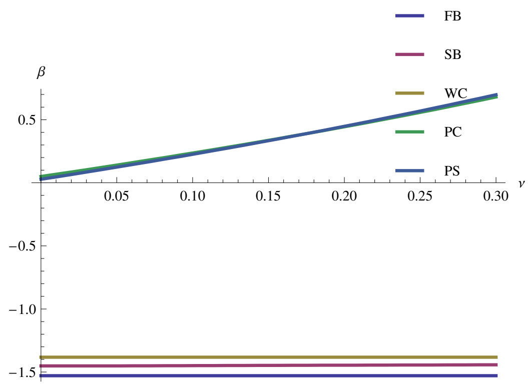
1.4.3 Sensitivity of β to Contract Power

In this section, we focus exclusively on the profit-sharing regime to assess the relationship between contract power and risk. Profit-sharing is the only system described in Section 1.3 in which the power of the contract can be explicitly obtained from the optimal policy. It is therefore the most natural choice to investigate the before mentioned relationship. In Figure 1.7, we plot the equilibrium values of risk (β) and power ($1-\tau$) implied in the profit-sharing policy against each of the five industry parameters, while keeping the remaining four at their representative levels ($d = 12, \psi = \tilde{\psi}(12), \sigma_\theta = 0.10, \nu = 0.15$ and $\rho = -0.25$).

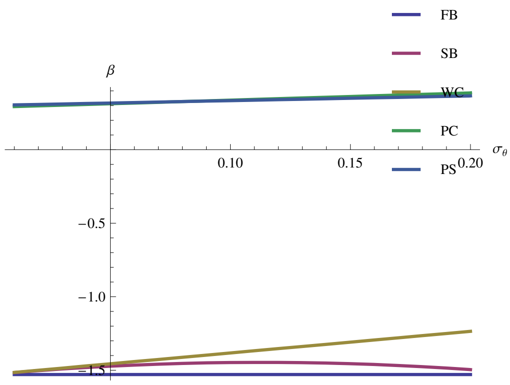
In which regards the cost-related variables, the simulation results suggest that they do not exert much influence on the optimal contract power, as τ^{ps} is fairly constant across σ_θ and ρ. The disutility of effort, however, has a non-negligible impact: as expected, as ψ increases, it becomes more expensive to induce the firm to reduce costs, so the contract power optimally decreases, as shown in Figure 1.7(e). But in this case, β_{PS} is constant with ψ, so we conclude that there is no significant relationship among the optimal levels of power and risk when cost conditions change.



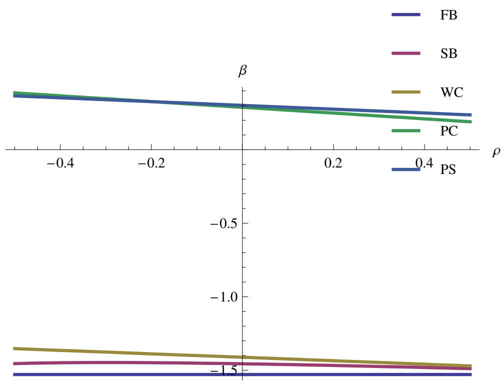
(a) Sensibility to d



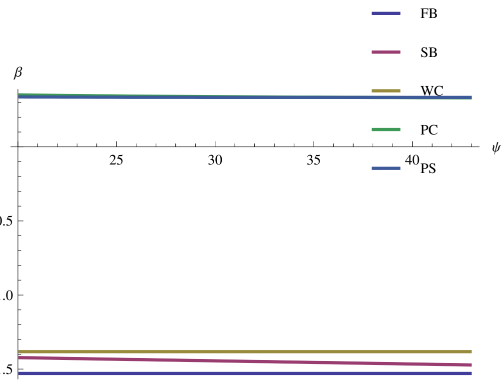
(b) Sensibility to ν



(c) Sensibility to σ_θ

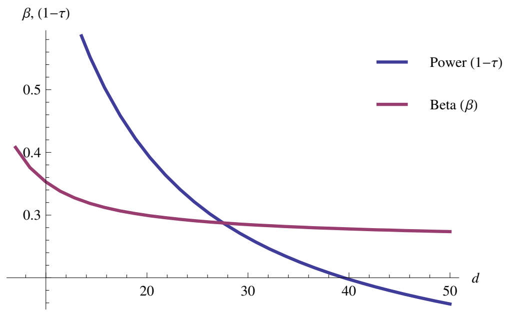


(d) Sensibility to ρ

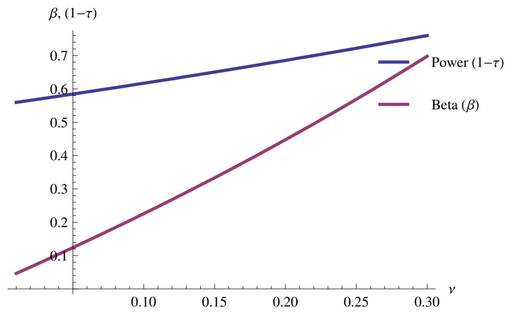


(e) Sensibility to ψ

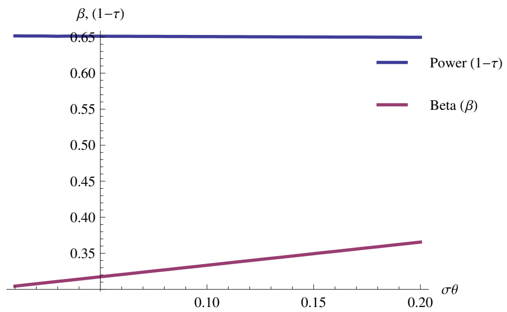
Figure 1.6: β profiles by regulatory regime.



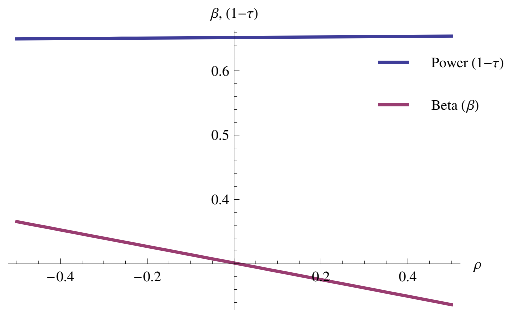
(a) Sensibility to d



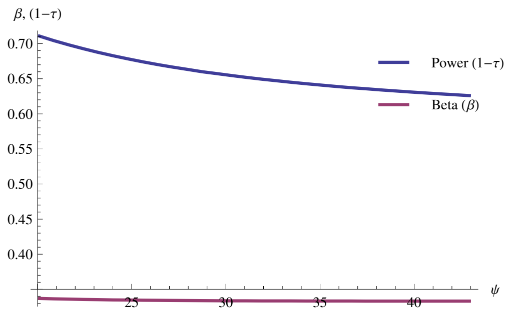
(b) Sensibility to ν



(c) Sensibility to σ_θ



(d) Sensibility to ρ



(e) Sensibility to ψ

Figure 1.7: Contract power and risk implied by PS regulation.

On the other hand, demand variables do play a significant role in shaping the profile of β and τ, and the numerical results suggest a positive relationship among power and risk. Figure 1.7(a) shows that the higher the elasticity of demand, a quantity indirectly measured by the intercept parameter d, the higher are both $(1-\tau^{ps})$ and β_{PS}. As a matter of fact, Figures 1.4(a), 1.4(c) and 1.4(e) combined have already shown that β_{PS} decreases with d, which is the same effect. And in Figure 1.7(b) we see that as demand uncertainty

rises, power and risk also increase, monotonically.

One additional feature worth of comment is that, in our simulated economy, the optimal power not only varies with demand conditions but it is also never greater than 80%, at least in the cases depicted in Figure 1.7. This implies that a pure price-cap mechanism, in which power is 100% whatever the market conditions or firm characteristics are, is rarely optimal.

1.5 Conclusions

We expanded, in a parsimonious way, the regulatory model of Laffont and Tirole (1986, 1993) to incorporate aggregate risk in the cash flows of the firm, and showed how the gradual introduction of restrictions in the space of contracts lead to an increase in the systematic risk of the firm. In incentive theory, from where the modern regulatory models were derived, the principal is usually allowed to operate direct transfers to the agent, but this ability prevents the model from generating sensible levels of risk to the firm. Only when the regulator is restrained from using this instrument we are able to reproduce CAPM betas like those observed for regulated companies in the real world.

By simulating our model, we managed to uncover the effect that major industry characteristics have on the risk of a regulated firm, and the results obtained were in line with the usual beliefs of how β is supposed to behave or respond to those characteristics.

We also investigated how contract power and risk are related, using a model of profit-sharing regulation in which both power and risk are endogenously determined. While the general evidence provided by the numerical simulations is not overwhelming, the results obtained are indicative that power and beta are positively related, and this result can be seen as an internally consistent formalization of the generally accepted view that powerful regulatory systems induce greater risk on the firm. This result corroborates previous empirical studies such as Grout and Zalewska (2006) and Alexander, Mayer and Weeds (1996). However, by providing a fully developed model of regulation that disentangles the effect of each variable on risk and power, our interpretation is that the empirical results obtained in those studies were more likely to be generated by demand instead of cost conditions.

Chapter 2

Regulation and Return Asymmetry

Regulated firms may be subject to some regulatory practices that potentially affect the symmetry of the distribution of their future profits. If these practices are anticipated by investors in the stock market, the pattern of asymmetry in the empirical distribution of stock returns may differ among regulated and non-regulated companies. In this paper We review some recently proposed asymmetry measures that are robust to the empirical regularities of return data and use them to investigate whether there are meaningful differences in the distribution of asymmetry between these two groups of companies

Keywords: Economic regulation, skewness, non-parametric methods.

JEL Classification: L51, G12

2.1 Introduction

Economic regulation is generally motivated on efficiency grounds, for instance as a response from a benevolent central planner to a market failure such as a natural monopoly. In those situations, a formally constituted regulatory institution, acting on behalf of the public interest, tries to emulate the conditions of a long-run competitive market equilibrium, in which industry participants earn a rate of return on their investment compatible with their cost of capital. There are certainly other reasons, not necessarily welfare enhancing or based on the public interest, that also gives rise to regulation. Stigler (1971) and Peltzman (1976) for instance suggested that regulatory institutions are the outcome of special interest groups that compete for political influence and the allocation of rents, in what has been called the capture theory of regulation.

In either case, the actions and policies of the regulator have a direct and profound impact on the profitability and risk of regulated entities. Peltzman argued that the regulatory process attenuates profit fluctuations, so the risk of the firm decreases when regulation becomes more stringent. This "buffering" hypothesis has been interpreted as a reduction in the variance of profits and possibly on the (CAPM) beta of the regulated firm. This hypothesis was tested by several authors, e.g. Norton (1985) and Davidson, Rangan and Rosenstein (1997), and it seems to be valid at least in sectors under low-powered regulation. High-powered (incentive-based) regulatory systems such as the price-cap seem to impose more risk on the firm, as claimed for instance by Alexander, Mayer and Weeds (1996) and Grout and Zalewska (2006) who find a positive relation between contract power (incentives for cost reduction) and beta.

While much effort has been dedicated to investigate the relationship among regulation and the variance of profits, we argue that regulatory actions and policies also have a non-negligible effect on the *asymmetry* of returns on regulated businesses. It is not difficult to identify situations in which the authority may be interested in systematically generating significant (positive or negative) returns for the regulated firm. If this is the case, as these events unfold and are recognized by investors in the secondary market, we would be able to detect departures from symmetry in the distribution of stock returns.

There are at least three different regulatory phenomena that can potentially affect the symmetry of the distribution of profits of regulated firms. First, the regulator may be captured by some interest group aligned with the industry under supervision, or by the government currently in office. The influence of these groups in the workings of the regulator is likely to generate positive returns for the regulated entities if the pressure group is aligned

with the industry, or negative returns otherwise. Examples of this sort of situation abound in the literature, the classic reference being cab regulation in New York City (Viscusi, Harrington and Vernon, 2005).

Second, regulated firms may be subject to expropriatory actions from the government, which vary from intentional delays in the rate revision process to a full-blown nationalization of assets without fair or timely compensation. These events will, most of the time, be accompanied by some sort of justification from the authorities trying to characterize them as legitimate actions taken in defense of the public interest. Regardless whether these events are indeed what the authorities claim them to be, the fact is that changes in investors expectations regarding the probability of such events are enough to cause some selling pressure on the firm's stock. This effect can become recurrent if the government fails to change those expectations, eventually altering the shape of the *ex-post* stock return distribution.

Third, in industries under the used-and-useful principle of valuation of the rate base, there is also a permanent downside risk for the firm every time the operating assets are scrutinized by the regulator. In the scheduled hearings for rate revision, some assets may be disallowed from the rate base when their acquisition is considered imprudent by the regulator, and no compensation will be due from such a discritionary action. While this practice is important to mitigate the well known Averch-Johnson effect of over-investment in rate-of-return regulated industries, their impact on the risk and in particular the symmetry of returns is clear. See for instance Kolbe, Tye and Myers (1993) for a richer treatment of this issue in the U.S. natural gas sector.

In any case, unless one can clearly identify which of the above mentioned phenomena is predominant in a given industry, which is certainly a very difficult task, it is generally not possible to determine, *ex-ante*, the combined effect of those actions on the asymmetry of returns of the regulated entities. But they all represent additional factors that affect the symmetry of returns and are shared only, in a lesser or greater extent, by regulated firms. In this paper we wish to test the hypothesis that regulated stocks exhibit returns that are "more asymmetric" than those of a control group formed by non-regulated stocks. To carry such a task we have to deal with the problem of (i) correctly measuring asymmetry in stock returns, and (ii) claiming that the cross-sectional distribution of return asymmetry differ for regulated and non-regulated stocks.

In order to formally state our hypothesis, consider the population of all traded stocks in a given market, and let $S \in \mathbb{R}$ denote some measure of the asymmetry in the distribution of realized returns. Let F_S denote the distribution of this asymmetry measure in the cross-section of stocks, such that $F_S(s)$ is the mass of companies with $S \leq s$, i.e. $N^{-1} \sum_{i=1}^{N} \mathbf{1}\{s_i \leq s\}$

where $\mathbf{1}\{\cdot\}$ is the indicator function. In the most general case, one could write the joint distribution $F_{S,\mathbf{X}}(s, x_1, \ldots, x_k)$ where X_j, $j = 1, \ldots, k$ are variables that might be related with S in the population. We concentrate on a simple case with a single explanatory variable. We write $F_{S,R}(s, r)$, where R is a categorical variable indicating whether the firm is regulated ($r = 1$) or not ($r = 0$). Finally, let $f_{S,R}(s, r)$, $f_S(s)$ and $f_R(r)$ denote the joint and marginal density functions associated with $F_{S,R}(s, r)$. In section 2.3.1 we investigate whether the distribution of asymmetry across firms is different among regulated and unregulated ones by testing the null

$$H_0 : f_{S,R}(s, r) = f_S(s) f_R(r), \tag{2.1}$$

for almost all $s \in \mathbb{R}$ and all $r \in \{0, 1\}$. As we do not want to impose any *a priori* structure on $f_{S,R}$, we use a number of non-parametric approaches to test (2.1) against the general alternative $H_A : f_{S,R}(s, r) \neq f_S(s) f_R(r)$ for some $s \in \mathbb{R}$ or $r \in \{0, 1\}$. The first is a test based on 2×2 contingency tables formed after an appropriate dichotomization of S. One advantage of this approach is that we can provide a corrected p-value that explicitly takes into account the fact that we are using estimates of asymmetry in lieu of actual observed data. The exact p-value under this situation will depend on the size and power of each asymmetry test, and the applied correction yields valid inference under measurement errors in S.

We also perform a test of (2.1) based on estimates of the conditional distribution $F_{S|R}(s \,|\, r)$, testing whether $H_0' : F_{S|R}(s \,|\, r) = F_S(s)$ for almost all $s \in \mathbb{R}$ and all $r \in \{0, 1\}$ against a general alternative $H_A' : F_{S|R}(s \,|\, r) \neq F_S(s)$ for some $s \in \mathbb{R}$ or $r \in \{0, 1\}$. As a matter of fact it is easier to test H_0' indirectly through the necessary condition

$$F_{S|R}(s \,|\, r = 0) = F_{S|R}(s \,|\, r = 1), \tag{2.2}$$

and to that end we use three different approaches. First we apply the Wald-Wolfowitz runs test of randomness on the sequence of zeros and ones in $\{r_i\}$ ordered by the corresponding asymmetry measure $\hat{s}_i$. Second, we use the Kolmogorov-Smirnov, Cramer-von Mises and Anderson-Darling tests for equality of distributions, which are based on the empirical CDF (i.e. non-smooth) estimates of the functions in (2.2). We use the improved versions of these tests, developed by Zhang (2006), which have greater power against any differences in the distributions, including location, scale and shape. Lastly, we use kernel smoothing methods to estimate the distributions in (2.2) and a distance measure among them to test whether they differ.

The hypothesized effect of regulation on return asymmetry can imply other statistical features on the conditional distribution $F_{S|R}(s \,|\, r)$ that can

be tested. If regulation increases the likelihood of asymmetric returns in either direction, then $F_{S|R}(s \,|\, r = 1)$ is expected to show a larger spread and we may test (2.2) against a scale alternative of the form

$$H'_S : F_{S|R}(s \,|\, r = 0) = F_{S|R}(\nu\, s \,|\, r = 1)\, \forall\, s \in \mathbb{R} \text{ and some } \nu \geq 1, \qquad (2.3)$$

a one-sided test that can be implemented with the Ansari-Bradley or the Mood statistics, for instance. On the other hand, if we consider any strictly positive measure $\mathcal{S}$ that does not discriminate among right or left asymmetry, then the density of $\mathcal{S}$ conditional on $R = 1$ is expected to be a positive location shift,

$$H'_L : F_{\mathcal{S}|R}(s \,|\, r = 0) = F_{\mathcal{S}|R}(s - \theta \,|\, r = 1)\, \forall\, s \in \mathbb{R}^+ \text{ and some } \theta \geq 0, \quad (2.4)$$

again a one-sided test. We use the Wilcoxon Rank-Sum statistic to test (2.2) against this specific alternative.

While we conjectured possible effects of economic regulation on return asymmetry, and formulated specific predictions that can be falsified by data, there are two major empirical challenges in testing the hypotheses (2.1) to (2.4). First, return asymmetry is not observed and must be estimated, so our fundamental piece of information is measured with error. In order to alleviate this problem, we rely on several nonparametric methods which are robust to the sampling error embedded in the asymmetry estimates.

Second, even if we were able to observe true asymmetry levels for each firm, we lack the necessary statistical design to isolate the effect of economic regulation from other factors in determining asymmetry. Ideally, we would like to have a sample of firms from various industries and "apply" regulation to a randomly selected subsample, obtaining control and treated groups. The ideal sample would have regulated and unregulated firms within in the same industry, and those subject to regulation are so because of an exogenous sampling process. This is obviously a practical impossibility, as regulation is the result of a comples political process and generally applies to all firms, in a larger or lesser extent, in the same industry.

Therefore, the extent to which we can argue whether regulation affects asymmetry is conditional to the sample used and no general causal relation can be established. We simply cannot have a group of comparable companies, operating in the same lines of businesses and with the same characteristics as the regulated ones but lacking only the regulatory restrictions, so that the confounding factors are controlled for. It is perfectly conceivable that some regulated industries have intrinsic characteristic that also impact return asymmetry, and if those characteristics are missing in the group of unregulated companies it would be hard to disentangle the impact exclusively due to

regulation. We try to alleviate this problem by considering several different regulated industries (water, oil, gas, energy, banks, transport and telecoms) in our sample, each with its own economics and specificities, hoping that their idiosyncrasies cancel each other out and their shared regulatory nature prevails. Fortunately, there are substantial differences in the economics of these regulated sectors, both in their cost and demand structures, so we believe that other asymmetry-affecting attributes they might have are (at least partially) controlled for by aggregation.

2.2 Measuring Asymmetry in Financial Returns

First we need to establish some notation and discuss the model of returns that will be used throughout the paper. Let y_t denote the (log) return of a given stock in period t and $\mathcal{F}_t$ the σ-algebra generated by the information available at time t. The traditional model for y_t in empirical applications is

$$y_t = \mu(\theta, \mathcal{F}_t) + \eta_t, \tag{2.5a}$$

$$\eta_t = h(\lambda, \mathcal{F}_t)\epsilon_t, \tag{2.5b}$$

where ϵ_t is an i.i.d. innovation with zero mean and unit variance, $\mu_t \equiv \mu(\theta, \mathcal{F}_t)$ and $h_t \equiv h(\lambda, \mathcal{F}_t)$ are the mean and standard deviation of y_t conditional on $\mathcal{F}_t$, θ and λ are vectors of parameters and $\eta_t \equiv y_t - \mu_t = h_t\epsilon_t$ is the (possibly heteroscedastic) residual. In Appendix B.1 we show the implications of model (2.5) for the moments of y_t, and how the conditional asymmetry of y_t is implied by the unconditional asymmetry of the innovations ϵ_t.

The concept of symmetry of a random variable can be formalized as follows:

Definition 1. *A real random variable Y with c.d.f. F_Y and support in $\mathbb{R}$ is said to have a center of symmetry in $\phi \in \mathbb{R}$ if $F_Y(\phi + a) = 1 - F_Y(\phi - a)$ for all $a \in \mathbb{R}$.*

An immediate consequence of this definition is that $F_Y(\phi) = 1/2$, so ϕ must also be the median of Y. The converse is obviously not necessarily true, as the median does not need to be the center of symmetry, which may not exist for a given random variable. In the context of financial data we are usually interested whether the return process $\{Y_t\}$ is symmetric with regards to its conditional (or unconditional) mean, and in this case the definition above is equivalent as saying that $Z_t \equiv Y_t - \mathbb{E}[Y_t|\mathcal{F}_t]$ has the same distribution as $-Z_t$, or $F_{Z_t}(z) = 1 - F_{Z_t}(-z)$ and, if the p.d.f. exists, $f_{Z_t}(z) = f_{Z_t}(-z)$ for all $z \in \mathcal{Y}$. This definition is fairly general and does not yield a single or

obvious index of the degree of asymmetry that Y_t might exhibit. As a result, several measures of asymmetry have been proposed. The aim of this section is to review some recent developments that specifically address the statistical features of financial returns. In particular, if a process $\{Y_t\}$ has finite third moment, the most popular measure of asymmetry is the *skewness* coefficient

$$S(Y_t) = \frac{\mathbb{E}[(Y_t - \mathbb{E}Y_t)^3]}{\mathbb{E}[(Y_t - \mathbb{E}Y_t)^2]^{3/2}}. \tag{2.6}$$

In a model such as (2.5), we show in Appendix B.1 how the conditional mean μ_t, the conditional variance h_t and the innovation process ϵ_t contribute to the unconditional skewness of Y_t:

$$S(Y_t) = \frac{\mathbb{E}[(\mu_t - \mathbb{E}\mu_t)^3] + 3\,\mathrm{cov}(\mu_t, \eta_t^2) + \mathbb{E}h_t^3\mathbb{E}\epsilon_t^3}{(\mathrm{var}\,\mu_t + \mathbb{E}h_t^2)^{3/2}}. \tag{2.7}$$

In (2.7) we see that even when the conditional mean has a negligible effect (as in the case of most stocks), the unconditional skewness of Y_t will depend on the skewness of the innovation process and on the conditional variance h_t. However, if in (2.6) we consider expectations conditioned on the information available at the previous period ($\mathcal{F}_{t-1}$), the conditional skewness of Y_t,

$$S_{t-1}(Y_t) = \frac{\mathbb{E}_{t-1}[(Y_t - \mathbb{E}_{t-1}[Y_t])^3]}{\mathbb{E}_{t-1}[(Y_t - \mathbb{E}_{t-1}[Y_t])^2]^{3/2}} = \frac{h_t^3\mathbb{E}_{t-1}\epsilon_t^3}{(h_t^2\mathbb{E}_{t-1}\epsilon_t^2)^{3/2}} = \mathbb{E}\epsilon_t^3, \tag{2.8}$$

is related only to the unconditional third moment of the innovation process. In this paper we implement methods for measuring both the unconditional and conditional asymmetry of returns, and (2.7) and (2.8) above shows that we need to estimate a fully specified (and yet flexible) ARMA-GARCH model for the returns of each stock in our sample in order to understand where possible differences in both concepts of asymmetry might be coming from. We give further details on our modeling strategy for returns at the end of this section.

Cramér (1946) shows that, for i.i.d. normal data, the estimator $\widehat{S}$ of (2.6) based on sample moments and under the null of symmetry is asymptotically normal. More specifically,

$$\sqrt{T}\,\widehat{S} \xrightarrow{d} \mathcal{N}(0, 6), \tag{2.9}$$

where T is the sample size. Despite its widespread use, $\widehat{S}$ has some serious drawbacks when applied to financial returns that are not always considered in empirical applications. First, it is a moment-based measure and some financial assets are so heavy-tailed that their third moments might not exist.

Second, the statistic is centered on the unconditional mean of the series, which may also not be the best choice of location parameter for all assets. Stocks, for example, tend to exhibit some variation in their conditional mean, and a perfectly (conditionally) symmetric return series may be mistakenly interpreted as a skewed one by (2.6).

Third, the asymptotic variance in (2.9) crucially depends on the i.i.d normality of the data generating process. To understand how departures from this requirement affect the null distribution of $\widehat{S}$, suppose the process $\{Y_t\}$ is Gaussian but serially correlated. Conditional on high past values of Y_t, if the process is positively correlated, future values of Y_t are more likely to be also high. The nominal size of the test is understated and it is likely to detect spurious positive skewness, resulting simply from serial dependence. Lomnicki (1961) shows that if $\{Y_t\}$ admits a moving-average representation, then $\sqrt{T}(\widehat{S} - S)$ is asymptotically normal with zero mean and variance $6\sum_{j=-\infty}^{+\infty} \rho_j^3$, where ρ_j is the autocorrelation of order j of Y_t.

When data is i.i.d. but non-Gaussian, the asymptotic variance of (2.9) is also incorrect. In particular, if the sample comes from a symmetric but leptokurtic distribution, a test based on (2.9) would also reject the null of symmetry too often. In this case, Stuart and Ord (1987) shows that $\sqrt{T}\,(\widehat{S} - S)$ is asymptotically normal with zero mean and variance

$$9 - 6m_4/m_2^2 + m_6/m_2^3, \tag{2.10}$$

where m_4/m_2^2 is the coefficient of kurtosis, or the standardized fourth moment. This asymptotic variance requires the existence of the sixth moment, which may be too restrictive in the context of financial returns, as documented by Jansen and de Vries (1991), Loretan and Phillips (1994) and Lima (1997). More recently, Bai and Ng (2005) derived the limiting distribution of $\widehat{S}$ in the general case of non-Gaussian and (weakly) dependent data. However, their skewness test applies only to series that are stationary up to the sixth-order, which is not the case of many financial returns that typically exhibit fat tails.

Lastly, despite being consistent, $\widehat{S}$ is biased in finite samples, simply because sample moments are not necessarily unbiased estimators of their population counterparts. Specifically, if $\widehat{m}_k$ is the k-th ordered sample moment, then $\mathbb{E}[\widehat{m}_2] = m_2(T-1)/T$ and $\mathbb{E}[\widehat{m}_3] = m_3(T-1)(T-2)/T^2$. There are simple small sample corrections for $\widehat{S}$, but they are not unique and different statistical packages usually employ different formulations with different finite-sample properties, as discussed by Joanes and Gill (1998). However, when the sample size is large this bias is negligible.

It is also known the estimation of higher-order moments is sensitive to

outliers, and the presence of a few extreme returns may lead an statistical test to indicate asymmetry in an otherwise symmetric distribution. In order to mitigate this effect robust measures of skewness have been suggested. Analogous to the role that the median and the interquartile range play as robust measures of location and scale, Hinkley (1975) considers a general class of asymmetry measures based on quantiles,

$$S_R(p) = \frac{(Q_{1-p} - Q_{0.5}) - (Q_{0.5} - Q_p)}{Q_{1-p} - Q_p},$$

where Q_p for $p \in (0, 1/2)$ denotes the p-th quantile, i.e. $F(Q_p) = p$ and $Q_{0.5}$ is the median. Bowley (1937) skewness statistic is a particular case in which quartiles ($p = 0.25$) are used, and

$$\sqrt{T}\,\widehat{S}_R(0.25) \xrightarrow{d} \mathcal{N}(0, 1.839) \tag{2.11}$$

in i.i.d. normal samples under the null of symmetry. However, the choice of $p = 0.125$ is also a valid statistic based on octiles and puts more weight on the tails of the distribution than Bowley's measure. It is not clear which p should be chosen for each application, and one could also integrate $S_R(p)$ over p to obtain $(\mathbb{E}Y - Q_{0.5})/\mathbb{E}|Y - Q_{0.5}|$, a skewness measure free of the p parameter and bounded on $(-1, 1)$. In spite of being robust to outliers, the limiting distributions of the above statistics, to the best of our knowledge, were derived only for the for the i.i.d. normal or exponential cases, by Moors et al (1996) and for the general i.i.d. case, by Ngatchou-Wandji (2006).

Neither (2.9), (2.11) or the Bai and Ng (2005) tests are adequate to detect asymmetry in the empirical distribution of stock returns, which has been shown to exhibit leptokurtosis, serial dependence and time-varying volatility, specially when calculated at short (daily or weekly) intervals (Tsay, 2005). Also, the possibility of arbitrage suggests that asymmetry, at least unconditionally, is not likely to be very substantial or persistent. Singleton and Wingender (1986) confirmed earlier evidence that *ex-post* common stock returns tend to be positively skewed in the cross-section, but showed that this phenomenon is not persistent for individual stocks or portfolios. Hence, better measures are required and in order to be valid and useful for financial data, symmetry tests must not only be robust to the stylized facts of stock returns but should also be powerful enough to detected small but significant departures from symmetry. Both of these desirable features are usually in conflict with one another, so it seems natural and appropriate to apply more than a single symmetry test as they vary in terms of robustness and power.

In the last few years a handful of symmetry tests were specially designed to be applied to financial data, and four of these tests are appropriate to

our study. Lisi (2007) uses the traditional skewness coefficient (2.6) applied to the residuals of a properly fitted ARMA model of returns, relying on bootstrapped standard errors for inference. The method of Chen and Lin (2008) explores the fact that for any process Z_t that is symmetric, mean-zero and stationary, $\mathbb{E}[\phi(Z_t)] = 0$ for any odd function $\phi(\cdot)$. The test of Bai and Ng (2001) for conditional symmetry is based on the empirical distribution of the innovations of a fully specified ARMA-GARCH model of returns. Finally, the Maasoumi and Racine (2009) test, which measures the distance between kernel density estimates of Y_t and its rotated series $\widetilde{Y}_t \equiv -Y_t + 2\mu_y$, where μ_y is the location parameter of Y_t. Under symmetry, Y_t and $\widetilde{Y}_t$ are equally distributed, so the distance between kernel estimates of their densities must be close to zero. The test can also be applied to consistent estimates of the innovations ϵ_t and $\widetilde{\epsilon}_t$, serving as a test of both conditional and unconditional symmetry of returns.

In order to perform these tests, we must specify and estimate a model such as (2.5) for each stock in our sample. We used an ARMA process for the mean equation and a APARCH specification for the variance equation, following Ding, Granger and Engle (1993),

$$h_t^{\delta} = \omega + \sum_{i=1}^{p} \alpha_i (|\eta_{t-i}| - \gamma_i \eta_{t-i})^{\delta} + \sum_{i=1}^{q} \beta_i h_{t-i}^{\delta}. \tag{2.12}$$

Several GARCH models are nested in (2.12), and in practice we set $p = q = 1$ and $\delta = 2$ for most stocks in our sample, which effectively reduces (2.12) to the GJR-GARCH(1,1) model of Glosten, Jagannathan and Runkle (1993). The APARCH model is a fairly general and yet parsimonious specification to capture the stylized facts of conditional variance in stocks returns (Teräsvirta, 2009). Most importantly, we assume that ϵ_t follows a skewed version of the Student-t distribution, in the sense of Fernandez and Steel (1998), which is a flexible specification that accommodates asymmetry and leptokurtosis in the data. Specifically, for each stock $i = 1, \ldots, N$ in our sample, we assume

$$\epsilon_{i,t} \sim \text{skewed-t}(\xi_i, \nu_i), \tag{2.13}$$

where parameters ν_i and ξ_i control the tail mass and asymmetry, respectively (see Appendix B.2 for further details on this distribution). This flexibility allows for consistent estimation of θ and λ under the possibility of ϵ_t being truly skewed. Newey and Steigerwald (1997) showed that if the conditional mean is not identically zero, the quasi-maximum likelihood estimator (QMLE) of conditional heteroscedasticity models such as (2.5) and (2.12) is consistent if both the true and assumed error density are symmetric. When this symmetry condition does not hold, either a flexible distribution (i.e., that allows

for skewness) must be assumed for ϵ_t or an additional parameter must be included in the mean equation so that the possibly non-zero location of the disturbance process can be identified. By considering a flexible distribution and estimating the mean and variance equations jointly, we preserve consistency of the QMLE of (2.5) and (2.12) under the possibility that the true innovations are skewed.

2.2.1 Unconditional Asymmetry Tests

Lisi (2007) addresses the problem of correctly quantifying the variance of $\widehat{S}$ by bootstrapping the residuals of a properly fitted ARMA model of returns. Using the notation in (2.5), a skewness coefficient is calculated for the filtered series, $\widehat{S}(\eta_t)$, and the variance of this statistic under the null of symmetry is obtained by resampling "symmetrized" residuals: if $\{\eta_1, \ldots, \eta_T\}$ is the original sample of residuals and $\eta_t^* \equiv |\eta_t - me(\eta_t)|$, where $me(\eta_t)$ is the median of η_t, then

$$\widetilde{\eta}_t = me(\eta_t) + z_t\,\eta_t^*, \quad t = 1, \ldots, T$$

where z_t is a Rademacher random variable, i.e., $P(z_t = -1) = P(z_t = +1) = 1/2$, is a reallocation of the original residual series that imposes symmetry. Resampling with replacement from $\{\eta_1, \ldots, \eta_T\}$ and calculating $\widehat{S}(\widetilde{\eta}_t)$ for each resampled series yield a bootstrapped distribution of $\widehat{S}(\eta_t)$ under the null of symmetry. This method does not exactly test the unconditional skewness of y_t but is an approximation that may work well when the conditional mean does not vary much, which seems to be the case for short-term stock returns. Lisi provides a simulation study showing that inference with the suggested resampling method has better size and power than with the asymptotic variance (2.10), with the advantage of requiring existence of the third moment of y_t only.

Chen and Lin (2008) proposed a family of symmetry tests for the unconditional asymmetry of a return series y_t that is robust to serial dependence and leptokurtosis. The test explore the fact that $\mathbb{E}[\phi(X)] \neq 0$ implies asymmetry of X for any odd function ϕ. For three different choices of odd functions, namely $\phi_s(z) = z^3$, $\phi_c(z) = z/(1+z^2)$ and $\phi_p(z) = \arctan(z)$, the authors develop two types of tests: the H test, based on the statistic

$$H_T = \frac{1}{T}\left[\sum_{t=1}^{T}\phi(z_t)\right]^{\top}\widehat{\Omega}_T^{-1}\left[\sum_{t=1}^{T}\phi(z_t)\right], \tag{2.14}$$

where $\widehat{\Omega}_T^{-1}$ is a HAC estimator of the long-run variance of $T^{-1/2}\sum_{t=1}^{T}\phi(z_t)$ and $z_t = (y_t - \bar{y})/\widehat{\sigma}_y$ is the standardized return; and the K test, based on

the statistic

$$K_T = \frac{1}{T}\left[\sum_{t=1}^{T}\phi(z_t)\right]^{\top}\widehat{C}_T^{-1}\left[\sum_{t=1}^{T}\phi(z_t)\right], \tag{2.15}$$

in which the long-run variance estimate $\widehat{\Omega}_T$ is replaced by a weighting matrix $\widehat{C}_T$ (see their paper for details). While $H_T \xrightarrow{d} \chi^2(1)$, the distribution of K_T is non-standard and its quantiles must be obtained by simulation: the 90% (95%) critical value is 28.31 (45.40). Applying each odd function ϕ_s, ϕ_c and ϕ_p to (2.14) or (2.15) result in six different test statistics: HS, HC, HP, KS, KC and KP. Note that in spite of being robust to the mentioned problems, these tests will only detect departures from symmetry and cannot discriminate against positive or negative asymmetry.

Maasoumi and Racine (2009) developed an entropy-based test of symmetry of a stationary (continuous or discrete) process $\{X_t\}$ based on the distance between the kernel density estimates $\widehat{f}_1$ of X_t and $\widehat{f}_2$ of $\widetilde{X}_t \equiv 2\mathbb{E}[X_t] - X_t$, i.e., a rotation of X_t about its mean. If X_t is symmetric, than $f_1(x) = f_2(\widetilde{x})$ almost surely, and the proposed test statistic $\widehat{S}_\rho$ is the integrated squared difference among the square-root of densities,

$$\widehat{S}_\rho = \frac{1}{2}\int_{-\infty}^{+\infty}[\widehat{f}_1^{1/2}(x) - \widehat{f}_2^{1/2}(\widetilde{x})]^2 dx. \tag{2.16}$$

Instead of developing asymptotic critical regions, the null distribution of $\widehat{S}_\rho$ is obtained by bootstrap: while the asymptotic critical values do not depend on the bandwidth of the kernel function (as this is a quantity that vanishes asymptotically), the value of the test statistic depends directly on the bandwidth. This is a major drawback for an asymptotic-based test, as variability of the test statistic across bandwidths may be substantial. The empirical distribution of $\widehat{S}_\rho$ under the null of symmetry is obtained by resampling with replacement from a pooled sample $Z = \{X_1, \ldots, X_T, \widetilde{X}_1, \ldots, \widetilde{X}_T\}$ of size $2T$, and calculating $\widehat{S}_\rho^*$ for each bootstrapped sample Z^*. With sufficiently many replications, say $\widehat{S}_{\rho,1}^*, \ldots, \hat{S}_{\rho,B}^*$, the p-value of the test statistic is the proportion of bootstrapped statistics $\widehat{S}_\rho^*$ that are at least as extreme as the value obtained from the original sample, $\sum_{b=1}^{B}\mathbf{1}\{\widehat{S}_{\rho,b}^* \geq \hat{S}_\rho\}/B$. As suggested by the authors, we used $B = 399$ replications and the stationary bootstrap of Politis and Romano (1994). Since the test statistic is a distance measure, it is always non-negative and can not distinguish between right or left asymmetry as well.

2.2.2 Conditional Asymmetry Tests

As the model (2.5), (2.12) and (2.13) is estimated for each stock $i = 1, \ldots, N$, a direct test of conditional asymmetry can be based on the statistical significance of the estimates $\widehat{\xi}_i$ of the innovation distribution. As discussed previously, the QML estimation under (2.13) is consistent and asymptotically normal, providing the distribution necessary for significance testing.

Bai and Ng (2001) developed a test of symmetry for the residuals ϵ_t for a model such as (2.5) that does not require stationary nor i.i.d. data. Their test is consistent, asymptotically distribution-free and requires only consistent estimates of the residuals, i.e., of the mean and variance equations of (2.5) in the sense of Newey and Steigerwald (1997). Their test is based on the empirical distribution functions of ϵ_t and $-\epsilon_t$, which under the null of symmetry have the same distribution, denoted here by $F(\epsilon)$. The building blocks of their CS statistic are as follows. Let

$$W_T(x) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} [\mathbf{1}\{\epsilon_t \leq x\} - \mathbf{1}\{-\epsilon_t \leq x\}] \tag{2.17}$$

denote the difference between the number of ϵ_t and the number of $-\epsilon_t$ that are less than or equal to x, then divided by $\sqrt{T}$. Under symmetry of ϵ_t, $W_T(x)$ should be small at all values of x, and if ϵ_t was observed for all t, then $\max_x |W_T(x)|$ would be a natural candidate for testing conditional symmetry. Since $\widehat{\epsilon}_t$ must be used, let $\widehat{W}_T$ denote the feasible estimator of (2.17) and for $x \leq 0$ define $S_T(x) = \widehat{W}_T(x) - \widehat{W}_T(0) + \int_x^0 h_T^-(y)dy$, where

$$h_T^-(y) = g_T(y) f_T(y) \left[\int_{-\infty}^{y} g_T(z)^2 f_T(z) dz \right]^{-1} \int_{-\infty}^{y} g_T(z) d\widehat{W}_T(z).$$

For $x > 0$, define $S_T(x) = \widehat{W}_T(x) - \widehat{W}_T(0) - \int_0^x h_T^+(y)dy$, where

$$h_T^+(y) = g_T(y) f_T(y) \left[\int_{y}^{+\infty} g_T(z)^2 f_T(z) dz \right]^{-1} \int_{y}^{\infty} g_T(z) d\widehat{W}_T(z)$$

and where f_T is a kernel estimator of the density $f(\epsilon_t)$ based on a sample of size T, and g_T is an estimator of $g = \dot{f}/f$. The proposed test is based on the statistic

$$CS = \max_x |S_T(x)| \xrightarrow{d} \max_{0 \leq s \leq 1} |B(s)|, \tag{2.18}$$

where $B(r)$ is a standard Brownian motion. The asymptotic critical values are obtained by simulation, and are 2.21 and 1.91 for the 5% and 10% significance levels, respectively.

Lastly, the Maasoumi and Racine $\widehat{S}_\rho$ test (2.16) described in section 2.2.1 can also be applied to the residuals ϵ_t for a test of conditional symmetry. The only necessary adjustment is on the resampling methodology for deriving the null distribution of the test statistic, which in this context can be the simple i.i.d. bootstrap.

2.2.3 Empirical Results of Asymmetry Tests

Our sample consists of daily (log) returns of Brazilian stocks traded at the Bovespa from July 1994 to July 2009, covering the first 15 years of low inflation after the Real stabilization plan. Excess returns are calculated over the 1-day interbank deposit certificate (CDI), a common proxy for the risk-free rate in Brazil. Only companies whose most liquid stock were negotiated in at least 70% of the trading sessions were selected. Less than 40 companies satisfy this criteria during the entire 15 years, so our sample would be extremely small and biased towards the survivors if selection were restricted to the full sample period. In order to mitigate this problem we divided the 1994-2009 period into five non-overlapping three year sub-periods, and for each sub-period we picked every company that satisfied the 70% presence criteria[1]. This subsampling strategy also helps to detect temporary departures of symmetry that might occur. We collected a total of 481 series of returns across all subperiods, with some firms contributing to more than one series.

We first applied the traditional and robust tests (2.9) and (2.11), and report in Table B.1 the number of firms with significant skewness at a 10% nominal level, by sub-period and regulatory status. As discussed, these tests are likely to have an incorrect size when applied to serially correlated and heavy-tailed data. The traditional, i.i.d. normal test for skewness rejects symmetry in almost 70% of cases, while the robust (quantile-based) test, also based on i.i.d. normal returns, rejects less than 30% of cases. Both measures however indicate a decrease in the relative number of asymmetric cases in more recent periods and a predominance of positive asymmetry.

The results of the conditional and unconditional tests described in Sections 2.2.1 and 2.2.2 applied to the same sample of stock returns are repro-

[1]Empirical studies with Brazilian stocks usually have to deal with the issue of small (but growing) liquidity. There are not too many companies quoted in every trading session and, to make matters worse, most firms have more than one equity class outstanding (usually two classes, one with and another without voting rights). The trade-off in lowering the liquidity threshold to increase the sample size (cross-sectional units) is the introduction of the statistical problems associated with non-synchronous or infrequent trading and missing data. We found that requiring presence in at least 70% of trading sessions generates samples with at least 100 different companies every year, each with no more than 5 consecutive days of missing data.

duced in Tables B.2, B.3 and B.4, for a significance levels of 5% and 10%. Out of the 481 return series considered, from 15% to 30% were considered unconditionally asymmetric at the 5% level by the various tests. Under a 10% significance, from 23% to 41% of cases are considered asymmetric, depending on the specific test, which shows how oversized the traditional test (2.9) can be when applied to non-normal, dependent data. Lisi's test, the only unconditional test capable of detecting left and right asymmetry, also shows that the vast majority of asymmetric returns are positively skewed. In conditional terms, asymmetry is also predominantly positive and appears to be more pervasive across firms and periods, as some 23% to 48% (at a 5% level) or 34% to 56% (at a 10% level) of innovation processes were rejected by the $\widehat{\xi}$, $\widehat{S}_\rho$ and CS tests.

2.3 Regulatory Status and Return Asymmetry

In this section we pursue four different approaches to test whether there is significant association among return asymmetry S and regulatory status R among the stocks in our sample, i.e., a test of hypothesis (2.1). It would be inappropriate to impose any *a priori* structure on $f_{S,R}$, the very object we want to study, so all methods employed to test whether S and R are related are nonparametric. Gibbons and Chakraborti (2003) and Li and Racine (2007) provide and extensive discussion of the tests considered in this section.

2.3.1 Contingency Tables

Hypothesis (2.1) states independence among return asymmetry S, a continuous variable not directly observable, and regulatory status R, a dichotomous variable indicating whether the firm is regulated or not. If S could be observed and discretized in some economically meaningful way, a contingency table could be formed and the relation between S and R could be tested. In particular, discretizing S into two categories would allow Fisher's exact test to be carried. For concreteness, let there be n_0 unregulated and n_1 regulated firms in our sample, and assume that the true asymmetry s_i of stock returns of firm $i = 1, \ldots, n_0 + n_1$ is observed. Let $s^* > 0$ be some cutoff point (determined in terms of economic significance, for instance) such that we classify as "symmetric" every return distribution i such that $|s_i| < s^*$, and "asymmetric" otherwise. Let k denote the total number of firms categorized as asymmetric, and among these k firms let x be the subset of regulated ones. These four values completely describe a contingency table such as Table B.5.

Under the null hypothesis of independence and with fixed marginals n_0, n_1 and k, Fisher (1922) derived the exact probability of observing x in the upper left cell,

$$P_H(x \mid n_0, n_1, k) = \binom{n_1}{x}\binom{n_0}{k-x}\bigg/\binom{n_0+n_1}{k} \tag{2.19}$$

which is the hypergeometric distribution. For 2×2 tables, Fisher's exact test for independence is equivalent to test whether the odds ratio,

$$OR(x) = \frac{x}{n_1 - x}\bigg/\frac{k-x}{n_0-(k-x)}, \tag{2.20}$$

is significantly different from 1, a two-sided hypothesis. Our conjecture is that a larger proportion of regulated companies tend to exhibit asymmetry, so the appropriate alternative is $OR > 1$, a one-sided test. The exact p-value of this one-sided test is obtained by adding the probabilities of each table arrangement that yield an odds ratio higher than (2.20), conditionally on the observed marginals. The only way to increase OR keeping the marginals fixed is to transfer observations from the cell (Regulated $\cap$ Symmetric) to (Regulated $\cap$ Asymmetric) and, at the same time, from (Unregulated $\cap$ Asymmetric) to (Unregulated $\cap$ Symmetric), i.e., by increasing x. This can be done until any of the "donating" cells becomes empty. Then, the one-sided p-value of Table B.5 for testing $OR(x) > 1$ is

$$p\text{-val} = \sum_{i=1}^{L} P_H(x+i \mid n_0, n_1, k), \tag{2.21}$$

where $L = \min\{n_1 - x, k - x\}$.

In practice however S must be estimated, and the sampling error in $\widehat{S}$ does not allow us to observe the true entries in Table B.5. Fisher's test applied on the actually observed table would inform an incorrect p-value due to possible misclassification. Fortunately, under some assumptions, we can work around this problem. Consider any estimator $\widehat{S}$ of S and a test of $S = 0$ with rejection probability α and power $1-\beta$. We expect the following misclassification: α% of the companies with symmetric returns will fall in the asymmetric bin, while β% of the companies with asymmetric returns will be mistaken for symmetric ones. If the test error probabilities are independent of R, and if the power of the test is constant under the alternative ($S \neq 0$), the cross tabulation actually observed would be given by the entries A, B, C and D shown in Table B.6, in which actual entries are rounded to the nearest integer.

Row totals $n_0 = A + B$ and $n_1 = C + D$ would still be observed without error. Given α and β, the correct p-value for the one-sided test $OR > 1$ under measurement error in S can be calculated by

$$p\text{-val}(\alpha, \beta) = \sum_{i=1}^{L(\alpha,\beta)} P_H \left(\frac{A - n_1\alpha}{1 - (\alpha + \beta)} + i \,\middle|\, n_0, n_1, \frac{A + C - (n_0 + n_1)\alpha}{1 - (\alpha + \beta)} \right), \tag{2.22}$$

where $L(\alpha, \beta) = \min \{(n_1(1 - \beta) - A), (C - n_0\alpha)\} / (1 - (\alpha + \beta))$.

For each symmetry test reviewed in Sections 2.2.1 and 2.2.2, we calculate (2.22) with $\alpha = 5\%$ and 10% and $\beta = 5\%, 20\%$ and 35% to allow for different sizes and powers these tests may have in our case. Tables B.7, B.8 and B.9 show the results of this sensitivity analysis. Each asymmetry test generates a contingency table such as Table B.6 with an associated odds ratio. In only a few cases there was evidence that the true odds ratio was greater than 1, specially during the first period (1994-1997) and for conditional asymmetry. However, the evidence for an odds ratio different than one (a two-sided test) was much stronger. Most of the cases in which the null was rejected the estimated odds ratio was less than one, which means that regulated stocks exhibited *less* skewed returns than their non-regulated counterparts. This phenomenon was significant in all 5 subperiods, specially regarding conditional asymmetry. Table B.10, which aggregates and summarizes the results presented in Tables B.7, B.8 and B.9, provides supporting evidence, which seems to be robust to different values of α and β.

A limitation of this approach is that it cannot be applied to every observed table. For any given pair of test error probabilities (α, β), there is a lower and an upper bound for the observed odds of being asymmetric (A/B or C/D), given by

$$\alpha/(1 - \alpha) \leq A/B \text{ or } C/D \leq (1 - \beta)/\beta, \tag{2.23}$$

beyond which the proposed transformation can not be applied as it would generate negative "true" entries. This is so because we are assuming that the symmetry tests have constant power across the entire region of asymmetry. As test power usually increases as we move further away from the null, and (2.22) does not take this into account, a fixed level of β tends to overestimate (underestimate) the true power of the asymmetry test for stocks that are closer to (further from) the null. In tables with extreme odds, beyond the bounds in (2.23), the transformation proposed fails to generate sensible values for the associated true classification.

2.3.2 Equality of the Conditional Densities

In this section we perform three types of tests to verify the validity of hypothesis (2.2). The first is based on the ranks of the R variable when the sample is ordered by the estimated asymmetry coefficients. The other two are based on estimates of the conditional distribution in (2.2). Non-smooth estimates of $F_{S|R}(s|r=0)$ and $F_{S|R}(s|r=1)$ are compared with a modified version of the traditional Kolmogorov-Smirnov, Cramér-von Mises and Anderson-Darling tests, while smooth (kernel-based) estimates are compared with the entropy metric of Maasoumi and Racine (2002).

Consider the results of an asymmetry statistic $\widehat{S}$ applied to a set of N stock returns, and sort the bivariate sample $\{(\widehat{s}_i, r_i), i = 1, \ldots, N\}$ according to the value of $\widehat{s}$. The resulting arrangement of R in the ordered sample

$$\{(\widehat{s}_{(1)}, r_{(1)}), \ldots, (\widehat{s}_{(N)}, r_{(N)})\}$$

can be tested for randomness, which is also a test of (2.2). Too few runs in the sequence of zeros and ones in $\{r_{(1)}, \ldots, r_{(N)}\}$ is evidence against the equality of the above conditional distributions. One advantage of this test is that it is robust to small perturbations in the values of $\widehat{s}_i$, as long as these perturbations do not change the ranks too much, so the problem of measurement error in S is alleviated.

Table B.11 shows the results of the runs test applied to each sequence of $\{r_{(i)}\}$ formed after ordering the estimates from the 11 tests of asymmetry considered. There is not enough evidence to reject (2.2) on the basis of this test: only for the 1997-2000 period there exists some indication of non-randomness in unconditional asymmetry, as 5 out of the 8 tests rejected the null.

Another strategy for testing (2.2) is to partition the sample $\{(\widehat{s}_i, r_i), i = 1, \ldots, N\}$ conditional on the value of R and to apply one of the classical tests for equality of distribution in the two samples

$$\{(\widehat{s}_1, 0), \ldots, (\widehat{s}_{n_0}, 0)\} \text{ and } \{(\widehat{s}_1, 1), \ldots, (\widehat{s}_{n_1}, 1)\}$$

obtained. We use more powerful versions of the classical tests of Kolmogorov-Smirnov (K-S), Cramér-von Mises (C-vM) and Anderson-Darling (A-D) proposed by Zhang (2006). Concretely, let $\widehat{F}_0$, $\widehat{F}_1$ and $\widehat{F}$ denote the empirical distribution function of the unregulated, regulated and combined samples, respectively. Instead of the usual chi-squared distance criteria, Zhang (2006) suggests the following likelihood-ratio statistic

$$G(s) = 2\sum_{j=0}^{1} n_j \left[\widehat{F}_j(s) \log \frac{\widehat{F}_j(s)}{\widehat{F}(s)} + (1 - \widehat{F}_j(s)) \log \frac{1 - \widehat{F}_j(s)}{1 - \widehat{F}(s)}\right] \tag{2.24}$$

as the distance metric for the K-S, C-vM and A-D tests, where n_0 and n_1 are the number of unregulated and regulated companies in the sample. The likelihood-ratio version of these statistics are, respectively,

$$Z_K = \sup_s |\, G(s) \,| \tag{2.25a}$$

$$Z_C = \int_{-\infty}^{\infty} G(s) d\widehat{F}(s) \tag{2.25b}$$

$$Z_A = \int_{-\infty}^{\infty} G(s) \widehat{F}^{-1}(s)[1 - \widehat{F}(s)]^{-1} d\widehat{F}(s) \tag{2.25c}$$

and their null distributions are obtained by enumerating all the possible and equally likely $n!/(n_0!n_1!)$ values of the statistic. A simulation exercise provided by Zhang (2006) showed that the the tests (2.25) are at least as powerful as their original versions for location differences, and much more powerful for changes in scale or shape. For our sample of stocks, obtaining an exact p-value is infeasible as $n_0 + n_1 \approx 100$, so an approximate p-value was obtained by Monte Carlo, randomly selecting 10,000 values from the null distribution.

Table B.12 reports the results of the Z_A test, as the Anderson-Darling is generally considered the most powerful among the three classical tests. The results obtained with Z_K and Z_C were similar and are available from the authors by request. The general conclusion is analogous to the results of the runs test: only during the 1997-2000 period there seemed to exist a significant difference in the distribution of asymmetry among regulated and non-regulated stocks, both in conditional and unconditional terms. While the evidence from the Z_A test for that period is strong, it does not indicate the source of discrepancy between $\widehat{F}_{S|R}(s|r = 0)$ and $\widehat{F}_{S|R}(s|r = 1)$, i.e., whether it is due to location, scale or shape differences.

Using the same partitioned sample it is possible to estimate the conditional density functions by kernel methods, calculate some metric of discrepancy among the estimated densities and decide whether the calculated distance is statistically significant. We performed a test based on kernel smoothing methods proposed by Maasoumi and Racine (2002) and Granger, Maasoumi and Racine (2004). Their distance metric, applied to our context, can be written as

$$\widehat{D} = \int_{-\infty}^{\infty} \left[\widehat{f}^{1/2}_{R|S}(s \,|\, r = 0) - \widehat{f}^{1/2}_{R|S}(s \,|\, r = 1) \right]^2 ds, \tag{2.26}$$

and the null distribution of $\widehat{D}$ is obtained by bootstrapping the statistic from the combined sample $\{\widehat{s}_1, \ldots, \widehat{s}_{n_0}, \widehat{s}_{n_0+1}, \ldots, \widehat{s}_{n_0+n_1}\}$. Table B.13 contains the

results of the $\widehat{D}$ test, and once again the period from 1997 to 2000 stands out as the only one with sufficient evidence of inequality between the conditional densities, for both unconditional and conditional asymmetry. The $\widehat{D}$ test is also incapable of providing a precise source for the discrepancy between the conditional densities.

In conclusion, we were not able to reject (2.2) consistently (i.e., for the majority of the asymmetry criteria and across all the sampled periods) by any of the three different techniques employed to measure the degree of discrepancy of the density of S conditional on the different regulatory statuses.

2.3.3 Location and Scale Differences

In Sections 2.3.1 to 2.3.2 we failed to reject the hypothesis (2.1) and (2.2) against the general alternatives H_A and H'_A, except for the period from 1997 to 2000. As we only employed robust, nonparametric techniques to test such hypothesis, this could be simply the result of the relative lower power these tests have when compared to their parametric counterparts. One possible way to improve inference is to develop more specific alternative hypothesis, motivated by the theoretical implications of our conjecture, which naturally lead to more powerful tests.

One expected outcome of our conjecture is that for any asymmetry measure $\mathcal{S} \in \mathbb{R}^+$ that does not discriminate among left or right skewness, the distribution of $\mathcal{S}$ among regulated stocks must have a larger location parameter than the distribution of $\mathcal{S}$ among unregulated firms. The Wilcoxon Rank-Sum test can be used to test whether there is a significant discrepancy among the locations of these distributions, in the form of hypotesis (2.4). The test statistic is $W = \sum_{i=1}^{N} i\, r_{(i)}$, where the sequence of 0's and 1's in $\{r_{(i)}\}$ is sorted according to the respective asymmetry estimates $\widehat{s}_i$ as used for the runs test in Section 2.3.2. While an exact null distribution can be calculated by enumeration when N is small, the normal approximation is accurate enough for combined samples of at least 12 observations, which is our case.

Table B.14 summarizes the results of the Wilcoxon Rank-Sum test of (2.4), for the various asymmetry concepts and sampled periods. The general conclusion from this test is somewhat similar and confirms the results from the previous tests: the period from 1997 to 2000 stands out as the only one in which evidence for a difference among the conditional densities is strong enough, and, as indicated by previous test, in the sense that returns of regulated stocks are less asymmetric on average. Table B.14 reports one-sided p-values based on the alternative hypothesis that the location parameter of $f_{\mathcal{S}|R}(s|r = 1)$ is larger than the location parameter of $f_{\mathcal{S}|R}(s|r = 0)$, i.e.,

$\theta \geq 0$. However, for the 1997-2000 period, the Wilcoxon test applied to 7 out of the 8 unconditional asymmetry measures reported p-values larger than 0.95, which means that the location parameter for the regulated group is smaller and statistically significant at 5% in a one-sided test in the opposite direction ($\theta \leq 0$). The same conclusion can be drawn regarding conditional asymmetry, as 2 out of the 3 metrics yield similar results for the 1997-2000 period.

One additional implication of our conjecture is that the distribution of any metric S that also informs the sign or direction of asymmetry is likely to have a larger variance among regulated companies. This formally translates to the hypothesis (2.3), which can be non-parametrically tested by the Ansari-Bradley or Mood statistics if we use the skewness metric of Lisi, for unconditional symmetry, or the ξ_i parameter of the GARCH innovation in (2.13), for unconditional symmetry[2]. Table B.15 reports the results of these tests, and shows that the larger variance effect might have happened during the 1994-1997 period in terms of unconditional asymmetry. The Mood test indicates the same also happened from 2000 to 2003. And regarding the 1997-2000 period, we refrain from drawing any further conclusions, as the location parameters of $f_{S|R}(s|r=0)$ and $f_{S|R}(s|r=1)$ are likely to be different (and unknown) and the distribution of the scale tests is not well defined in this case.

2.4 Conclusions

We conjectured that regulated companies are subject to certain phenomena absent in a unregulated environment, potentially affecting the symmetry stock returns. Using a sample of Brazilian publicly traded firms, we could not find consistent differences in the distribution of cross-sectional asymmetry coefficients among regulated and non-regulated firms, i.e., across several periods and different measures of asymmetry. There is evidence, however, that during the period from 1997 to 2000, non-regulated companies exhibited *less* asymmetric returns.

The period from 1997 to 2000 marks the emergence and early development stages of economic regulation of key industries in Brazil, and their transition from state-owned monopolies to private, regulated businesses. The federal agencies in charge of the electricity, telecommunications and oil sec-

[2]It is statistically valid to test for scale discrepancies after the location alternative has been tested, as these tests are asymptotically independent when the set of weights of the linear rank test statistic obeys some usual conditions, namely symmetry about their mean. See Anderson (2001) for a proof.

tors, ANEEL, ANATEL and ANP respectively, were created in the period from 1996 to 1998, in conjunction with other major reforms such as privatization and liberalization. Other industries such as gas, water and transportation also witnessed the appearance of state level regulation during the same period. It is likely that such significant changes in the environment of regulated businesses had an impact on the empirical distribution of stock returns in that period. Even under such radical transformations, we found evidence that regulated firms exhibited less skewed returns than their unregulated counterparts. The mechanism by which the establishment of large regulatory bodies affect the shape of ex-post stock returns is unclear and would be an interesting topic for further investigation.

The empirical results in this paper are subject to limitations of practical nature. First, due to the size of the Brazilian stock market, the number of companies and industries surveyed is relatively small and might not allow for sufficient variation in return asymmetry (which, after all, is not likely to be substantial or persistent) to appear. Second, we had to deal with the fact that asymmetry we cannot be observed but must be estimated. We tried to mitigate the measurement error problem by using only robust nonparametric tests, specially those based on ranks, which are unlikely to vary much as a result of sampling error. Another approach to alleviate this problem would be a Monte Carlo sensitivity analysis on the results of each test performed in Section 2.3. By randomly drawing a new, cross-sectional set of estimates from the sampling distribution of the asymmetry statistics, each test performed in section 2.3 could be repeated a large number of times, and the relative impact of sampling error could be assessed. This additional exercise would be important if we were able to reject the null hypotheses (2.1) and (2.2) with the primary set of estimated asymmetry coefficients, as a robustness check. However, since the data were not able to generate sufficient evidence against the null, this additional effort seems unnecessary.

Chapter 3

The Pricing of Coskewness Risk

Three different approaches to the capital asset pricing model of Kraus and Litzenberger (1976) are tested with recent brazilian data and estimated using the generalized method of moments (GMM) as a unifying procedure. We find that ex-post stock returns generally exhibit statistically significant coskewness with the market portfolio, and hence are sensitive to squared market returns. However, while the theoretical ground for the preference for skewness is well established and fairly intuitive, we did not find supporting evidence that investors require a premium for supporting this risk factor in Brazil.

Keywords: Higher moment CAPM, asset pricing, coskewness.

JEL Classification: G12

3.1 Introduction

In addressing the poor performance of the CAPM to explain the cross-sectional variation of returns, empirical research in asset pricing have explored multiple avenues. At least since Samuelson (1970), Jean (1971, 1973) and Rubinstein (1973), one direction has been to investigate higher moments of the aggregate wealth portfolio and its relation to the preferences of a representative investor. In many cases, asset returns have distributional properties that cannot be adequately described by the normal paradigm, and some of these non-normalities (such as skewness and kurtosis) are connected to the returns of the wealth portfolio in terms of co-moments (or systematic moments). Also, rational investors can be expected to enjoy odd moments of the return distribution, and to dislike the even ones. Hence, a pricing model that extends the single-factor (covariance) benchmark of Sharpe-Lintner to explicitly account for these statistical features seems a valid attempt to incorporate the empirical regularities of returns in a preference-based theory for the demand of risky assets.

Kraus and Litzenberger (1976) provided the first empirical test of the three-moment CAPM for the U.S. economy. In their model, a representative investor has a preference relation over the first three moments of the distribution of his uncertain end-of-period wealth W, and in market equilibrium expected returns are given by[1]

$$\bar{R}_i - R_f = \beta_i \frac{d\bar{W}}{d\sigma_W}\sigma_m + \gamma_i \frac{d\bar{W}}{dS_W} S_m, \tag{3.1}$$

where R_i and R_f are the returns on asset i and on the risk-free asset; $\beta_i = \mathrm{cov}(R_i, R_m)/\sigma_m^2$ and $\gamma_i = \mathbb{E}[(R_i - \bar{R}_i)(R_m - \bar{R}_m)^2]/S_m^3$ are standardized measures of the covariance and coskewness of R_i with respect to the return on the market portfolio R_m; $\bar{W}$, σ_W and S_W are mean, standard deviation and cubic root of the third central moment of W; σ_m^2 and S_m are the variance and cubic root of the third central moment of R_m, and the bar indicates expected value. We can interpret $b_1 \equiv [d\bar{W}/d\sigma_W]\sigma_m$ and $b_2 \equiv [d\bar{W}/dS_W]S_m$ as the market prices of beta and gamma risk, respectively. When investors care

[1] The equilibrium relation in (3.1) holds when the return on the market portfolio has a skewed distribution. When $\mathbb{E}[(R_m - \bar{R}_m)^3] = 0$, the standardized measure of coskewness γ_i does not exist and the equilibrium relation (3.1) becomes

$$\bar{R}_i - R_f = \beta_i \frac{d\bar{W}}{d\sigma_W}\sigma_m + S_{imm} \frac{d\bar{W}}{dS_W},$$

where $S_{imm} \equiv \mathbb{E}[(R_i - \bar{R}_i)(R_m - \bar{R}_m)^2]$ is the probabilistic definition of coskewness of R_i with respect to the market return R_m.

about the asymmetry of returns, in particular exhibiting non-increasing absolute risk aversion, they have a preference for (positive) skewness and are willing to trade-off expected return for skewness, so $d\bar{W}/dS_W < 0$. Hence, b_2 and S_m must have opposite signs: when the market portfolio is positively skewed ($S_m > 0$), b_2 is negative and assets with positive coskewness ($\gamma_i > 0$) are preferred, as this implies $\mathbb{E}[(R_i - \bar{R}_i)(R_m - \bar{R}_m)^2] > 0$ which means that R_i is more likely to be positive than negative when R_m varies. The opposite rationale applies when $S_m < 0$, as $\gamma_i > 0$ implies that $\mathbb{E}[(R_i - \bar{R}_i)(R_m - \bar{R}_m)^2] < 0$. Kraus and Litzenberger (KL) tested the following empirical version of (3.1),

$$(\bar{R}_i - R_f)/R_f = b_0 + b_1\hat{\beta}_i + b_2\hat{\gamma}_i + u_i, \tag{3.2}$$

using the two-stage methodology of Black, Jensen and Scholes (1972) and Fama and MacBeth (1973) on a sample of stock returns from 1936 to 1970, grouped into 20 equally-weighted portfolios formed by beta and gamma deciles. The results obtained were consistent with the hypotheses of their three-moment CAPM, i.e., $b_0 = 0$, $b_1 > 0$ and $b_2 < 0$, as the market return was positively skewed during the period. Later, Friend and Westerfield (1980) expanded the KL study by including other asset classes (bonds), using value and equally weighted market portfolios and distinguishing among different subperiods and types of markets. They found some, but not conclusive, evidence that investors may pay a premium for positive skewness in their portfolios.

Barone-Adesi (1985) attributes the differing conclusions of these early studies to the econometric methodology employed. Both KL and Friend and Westerfield (FW) use the traditional two-pass methodology, which introduces the errors-in-variables problem in the second-stage (cross-sectional) regressions. Inference is complicated even further by the high collinearity between estimated betas and gammas in both studies. Alternatively, Barone-Adesi stipulates a quadratic market model as the data generating process of returns, which is slightly different from the covariance-coskewness preference model of KL, and uses the arbitrage pricing theory (APT) as the equilibrium model to obtain parameter restrictions. Using both the methodology of Black, Jensen and Scholes and the multivariate approach of Gibbons (1980), supporting evidence for the KL hypothesis is found.

Sears and Wei (1985) investigate the properties of risk premia in the KL model, showing that the two-fund separation theorem assumed to arrive at the equilibrium condition (3.1) implies that the market risk premium and an elasticity coefficient proportional to the marginal rate of substitution between skewness and risk enter both b_1 and b_2. They show that (3.1) can be rewritten

as

$$\bar{R}_i - R_f = \beta_i \frac{1}{1+K_3}(\bar{R}_m - R_f) + \gamma_i \frac{K_3}{1+K_3}(\bar{R}_m - R_f). \qquad (3.3)$$

Hence the market risk premium $\bar{R}_m - R_f$ is actually measured by $b_1 + b_2$, and $K_3 = b_2/b_1 = [(d\bar{W}/dS_W)/(d\bar{W}/d\sigma_W)](S_m/\sigma_m)$. In a later study, Sears and Wei (1988) provide an empirical tests of the three-moment CAPM as formulated in (3.3) and find that K_3 is significant and has the correct theoretical sign.

Lim (1989) was the first test of (3.1) using the generalized method of moments (GMM), and was based on the specification of Sears and Wei in (3.3). The GMM framework allows for the simultaneous estimation of factor loadings and risk prices, avoiding the problems of the two-pass methodology and strong assumptions about the distribution of returns. While the estimated elasticity term K_3 was significant, $b_2 = 0$ could not be rejected. Lee, Moy and Lee (1996) estimated (3.3) using the multivariate approach of Gibbons (1992) and arrived at similar conclusions.

Barone-Adesi, Gagliardini and Urga (2004) followed the same general approach of Barone-Adesi (1985), namely a quadratic market model with APT derived restrictions, but used pseudo maximum-likelihood to estimate model parameters, a method equivalent to the iterated non-linear SUR of Gallant (1975). Using data from 1963 to 2000 and size-sorted portfolios, the authors found strongly significant sensitivities to market skewness, but an insignificant risk premium for coskewness risk.

All the empirical tests of the three-moment CAPM reviewed above are unconditional in nature, i.e., they do not allow for time variation in the parameters in an explicit way. Instead, most of these studies test the model in several subperiods (usually five years) to allow for potential shifts in risk prices and factor loadings. Some research on the role of *conditional* systematic skewness in asset pricing appeared recently. Harvey and Siddique (2000) find that conditional skewness is significant in explaining the cross-section of returns, even when the traditional Fama-French size and price-to-book factors are included. They also find conditional skewness to be an economically important feature of asset returns, commanding a premium of 3.60% per year on average. Smith (2007) corroborates this evidence with GMM estimates of a stochastic discount factor (SDF) compatible with a conditional three-moment CAPM, as opposed to full-information maximum likelihood estimation of the beta representation used by Harvey and Siddique. Dittmar (2002) uses the same GMM/SDF approach to estimate a four-moment conditional CAPM, that also includes time-varying co-kurtosis, but finds necessary to include the return on human capital (labor income) to arrive at an admissible pricing kernel and to drive out the importance of the Fama-French

factors.

Research on the validity of a three-moment CAPM in Brazil is still in its infancy. The major challenge is the lack of an adequately sized sample of assets returns, both in the cross-sectional as well as in the time dimension. Athayde and Flôres (2000) employ a GMM methodology analogous to the approach in Lim (1989), but use factor-mimicking portfolios to eliminate the preference parameters in (3.1) and arrive at a testable model consisting only of observed variables. Using daily returns of the 10 most liquid stocks in the São Paulo Stock Exchange during 1996-1997, the authors find that the inclusion of skewness, but not of kurtosis, brings a significant gain to the model (in terms of the J test), vis-à-vis the traditional two-moment CAPM. da Silva (2004) estimates a market model with higher moments using monthly returns of brazilian stocks from 1990 to 2003, which includes periods of high (pre 1994) and low inflation (post 1994). Stocks were sorted into 5 size and 5 price-to-book portfolios, and models with even higher moments (up to the tenth) and with the SMB and HML Fama-French factors were tested. Only time-series regressions were run, producing estimates of factor loadings but not risk premia, and coskewness was found not to be important when cokurtosis or the Fama-Franch factors were included. More recently, Castro Jr. (2008) worked with weekly returns of individual stocks for the 2003-2007 period, and using the two-pass methodology did not find coskewness to be priced.

There is room for further investigation of the adequacy of a higher-moment CAPM in the brazilian capital market, and the present study intends to contribute by applying a number of methodological improvements. First, the joint estimation of factor loadings and risk premia has not yet been attempted. Multivariate nonlinear methods can be implemented at a low computational cost, and eliminate the problems associated with the two-stage procedure. Second, due to the mixed results found in the earlier literature, it seems necessary to conduct a more thorough analysis, employing different sets of moment conditions across different asset aggregation criteria (equally and value weighted portfolios sorted by size, price-to-book and industry) and return horizons (weekly and monthly). Third, we can benefit from the additional degrees of freedom introduced by having a larger (but still not very large) dataset. The period from 2000 to 2009 provides 10 years of returns during which Brazil experienced the same, stable macroeconomic policy of floating exchange rates and inflation targeting. Also, the number of stocks and the average daily volume of transactions in the local stock exchange experienced an enormous growth since 2004, and this richer dataset has not yet been put to test the empirical implications of higher-moment asset pricing models. While the time dimension still seems too small for a fully conditional

model to be tested, an unconditional version with fixed coefficients could be estimated for different subperiods, e.g. 2000-2004 and 2005-2009, to allow for a possible change in parameters in the more recent and active stock market.

3.2 Skewness in Asset Pricing

Consider a set of N financial assets and let $\mathbf{R}$ denote the $(N \times 1)$ return vector, with mean $\boldsymbol{\mu} = \mathbb{E}[\mathbf{R}]$ and covariance matrix $\boldsymbol{\Sigma} = \mathbb{E}[(\mathbf{R}-\boldsymbol{\mu})(\mathbf{R}-\boldsymbol{\mu})']$. The coskewness of $\mathbf{R}$ is the third central moment, a $(N \times N \times N)$ tensor with typical element $S_{ijk} = \mathbb{E}[(R_i - \mu_i)(R_j - \mu_j)(R_k - \mu_k)]$. Following Athayde and Flôres (2004), we can conveniently write coskewness as the $(N \times N^2)$ matrix

$$\mathbf{S} = \mathbb{E}[(\mathbf{R}-\boldsymbol{\mu})(\mathbf{R}-\boldsymbol{\mu})' \otimes (\mathbf{R}-\boldsymbol{\mu})']. \tag{3.4}$$

Notice that S_{iii} is simply the third central moment of R_i, and that there are only $\binom{n+2}{3}$ non-identical elements in $\mathbf{S}$. If R_p is the return of a portfolio with asset weights given by the $(N \times 1)$ vector ω, the first three moments of R_p are $\omega'\boldsymbol{\mu}$, $\omega'\boldsymbol{\Sigma}\omega$ and

$$\omega'\mathbf{S}(\omega \otimes \omega) = \mathbb{E}\left[\sum_{\mathbf{i=1}}^{\mathbf{N}} \omega_{\mathbf{i}}(\mathbf{R_i} - \mu_{\mathbf{i}})(\mathbf{R_p} - \mu_{\mathbf{p}})^{\mathbf{2}}\right] = \omega'\mathbf{S_p} \tag{3.5}$$

where $\mathbf{S_p} = \mathbf{S}(\omega \otimes \omega)$ is the $(N \times 1)$ vector of coskewnesses between R_i and R_p. The contribution of asset i to the skewness of portfolio p is given by $S_{ipp} \equiv \mathbb{E}[(R_i - \mu_i)(R_p - \mu_p)^2]$. When investors have a preference for (positive) skewness, assets with $S_{ipp} > 0$ are preferred as volatility in R_p increase the likelihood of $R_i > 0$.

In asset pricing tests, the relevant measure of an asset's coskewness is against the return on the aggregate market portfolio R_m, or S_{imm}. This measure of *systematic* skewness is usually standardized by the third moment of the market portfolio, yielding $\gamma_i = S_{imm}/S_m^3$ à la Kraus and Litzenberger, where we define $S_m^3 \equiv S_{mmm}$ to align notation with (3.1). Interpretation of the gamma coefficient now depends on the sign of S_m^3, as explained in the introduction.

The inclusion of coskewness risk in asset pricing has been motivated on three different perspectives. The first one is a fully specified equilibrium model of portfolio formation[2], in which agents have preferences for moments of the distribution of their wealth and maximize expected utility,

[2]Consider a representative agent who has utility $U(W)$ over his uncertain end-of-period wealth W. A third-order Taylor expansion of $U(W)$ over $\bar{W} = \mathbb{E}W$ is $U(W) \approx U(\bar{W}) + U'(\bar{W})(W - \bar{W}) + U''(\bar{W})(W - \bar{W})^2/2! + U'''(\bar{W})(W - \bar{W})^3/3!$. Taking the expectation,

along the lines of Kraus and Litzenberger. The second is based on the APT, in which a statistical model for the return generating process is specified and testable restrictions are obtained from arbitrage arguments, as in Barone-Adesi, Gagliardini and Urga (2004). The third is based on direct estimation of the pricing kernel, which is modeled as a second-order expansion of an SDF that is a function of the returns on the aggregate wealth portfolio, as in Dittmar (2002). In the following sections we review each of these approaches and briefly explain the econometric methodology used for their estimation and testing.

3.2.1 A Fully Specified Equilibrium Model

Kraus and Litzenberger specified detailed conditions on preferences, beliefs and investor behavior to arrive at the equilibrium pricing equation (3.1). Accounting for the details raised by Sears and Wei (1985), and writing $K_3 = \phi(S_m/\sigma_m)$ where $\phi = (d\bar{W}/dS_W)/(d\bar{W}/d\sigma_W)$, we obtain the following empirical version of the KL model:

$$r_{i,t} = \left[\gamma_{1,i}\left(\frac{\sigma_m}{\phi S_m + \sigma_m}\right) + \gamma_{2,i}\left(\frac{\phi S_m}{\phi S_m + \sigma_m}\right)\right] r_{m,t} + u_{i,t} \qquad (3.7)$$

for each asset $i = 1, \ldots, N$, where $r_{i,t} = R_{i,t} - R_{f,t}$ and $r_{m,t} = R_{m,t} - R_{f,t}$ are the excess rates of return and $\mathbb{E}[u_{i,t}|r_{m,t}] = 0$ for all t. We can estimate and test the SW/KL model (3.7) under a generalized method of moments

---

we have

$$\mathbb{E}U(W) \approx U(\bar{W}) + \frac{U''(\bar{W})}{2!}\mathbb{E}[(W-\bar{W})^2] + \frac{U'''(\bar{W})}{3!}\mathbb{E}[(W-\bar{W})^3], \qquad (3.6)$$

which makes explicit how the third moment of W enters expected utility. Also, (3.6) satisfies the desirable properties listed by Arrow (1971) if $U' > 0$ (positive marginal utility), $U'' < 0$ (decreasing marginal utility, or risk-aversion), and $d[-U''/U']/dW \leq 0$ (non-increasing absolute risk aversion). It is easy to see that this last condition,

$$\frac{d}{dW}\left[\frac{-U''}{U'}\right] = \frac{-U'\,U''' + (U'')^2}{(U')^2} \leq 0$$

implies $U''' \geq (U'')^2/U'$, a strictly positive lower bound for the third derivative. Non-increasing ARA is therefore sufficient for $U''' > 0$ implying preference for (positive) skewness. Under such conditions, and invoking a separation theorem, equation (3.1) holds.

framework, using the following moment conditions:

$$\mathbb{E}\left[(\phi S_m + \sigma_m)r_i - (\gamma_{1,i}\sigma_m + \gamma_{2,i}\phi S_m)\, r_m\right] = 0, \tag{3.8a}$$
$$\mathbb{E}\left[(\phi S_m + \sigma_m)r_i r_m - (\gamma_{1,i}\sigma_m + \gamma_{2,i}\phi S_m)\, r_m^2\right] = 0, \tag{3.8b}$$
$$\mathbb{E}[(r_i - \mu_i)(r_m - \mu_m) - \gamma_{1,i}\sigma_m^2] = 0, \tag{3.8c}$$
$$\mathbb{E}[(r_i - \mu_i)(r_m - \mu_m)^2 - \gamma_{2,i}S_m^3] = 0, \tag{3.8d}$$
$$\mathbb{E}[r_i - \mu_i] = 0, \tag{3.8e}$$
$$\mathbb{E}[r_m - \mu_m] = 0, \tag{3.8f}$$
$$\mathbb{E}[(r_m - \mu_m)^2 - \sigma_m^2] = 0, \tag{3.8g}$$
$$\mathbb{E}[(r_m - \mu_m)^3 - S_m^3] = 0. \tag{3.8h}$$

Conditions (3.8a) and (3.8b) come directly from the specification of (3.7), while (3.8c) and (3.8d) come from the definitions of $\gamma_{1,i}$ and $\gamma_{2,i}$, respectively. There are a total of $5N + 3$ moment conditions[3], and $3N + 4$ parameters to be estimated ($\gamma_{1,i}$, $\gamma_{2,i}$ and μ_i for $i = 1, \ldots, N$, and ϕ, μ_m, σ_m and S_m), leaving $2N - 1$ overidentifying restrictions to be tested by Hansen's J statistic. The KL three-moment CAPM predicts $\boldsymbol{\gamma}_1 = (\gamma_{1,1}, \ldots, \gamma_{1,N})'$ and $\boldsymbol{\gamma}_2 = (\gamma_{2,1}, \ldots, \gamma_{2,N})'$ to be nonzero, and $\phi < 0$. The premia related to covariance and coskewness risks are estimated as

$$\hat{b}_1 = \frac{\hat{\sigma}_m \hat{\mu}_m}{\hat{\phi}\hat{S}_m + \hat{\sigma}_m}, \qquad \hat{b}_2 = \frac{\hat{\phi}\hat{S}_m \hat{\mu}_m}{\hat{\phi}\hat{S}_m + \hat{\sigma}_m}. \tag{3.9}$$

We expect $b_1 > 0$, and b_2 and S_m to have opposite signs.

3.2.2 Arbitrage Pricing

Under the linear factor model (LFM), excess returns have a factor structure of the form

$$r_{i,t} = \beta_{0,i} + \boldsymbol{\beta}_i' \boldsymbol{f}_t + \epsilon_{i,t} \tag{3.10}$$

[3] To facilitate the numerical optimization process, we linearized parameter cross-products in (3.8a) and (3.8b) using a first-order approximation around consistent estimates. We used

$$\gamma_{1,i}\sigma_m \approx \hat{\gamma}_{1,i}\sigma_m + \gamma_{1,i}\hat{\sigma}_m - \hat{\gamma}_{1,i}\hat{\sigma}_m,$$
$$\phi S_m \approx \hat{\phi} S_m + \phi \hat{S}_m - \hat{\phi}\hat{S}_m,$$
$$\phi S_m \gamma_{2,i} \approx \hat{\phi}\hat{S}_m \gamma_{2,i} + \hat{\phi} S_m \hat{\gamma}_{2,i} + \phi \hat{S}_m \hat{\gamma}_{2,i} - 2\hat{\phi}\hat{S}_m\hat{\gamma}_{2,i},$$

where a hat denotes a consistent estimate of the corresponding parameter. This procedure yields estimators that are asymptotically equivalent to the unmodified ones, see Fuller (1976), pp. 215-218.

for every asset $i = 1, \ldots, N$ and period t, where $\boldsymbol{f}_t = (f_{1,t}, \ldots, f_{K,t})'$ is the $K \times 1$ vector of factor realizations on date t, $\boldsymbol{\beta}_i = (\beta_{1,i}, \ldots, \beta_{K,i})'$ is the $K \times 1$ vector of factor loadings (sensitivities) of asset i with respect to the $K < N$ factors. Most importantly, the K factors (and the constant) are assumed to be orthogonal to the contemporaneous error term,

$$\mathbb{E}[\epsilon_{i,t} \mid f_{j,t}] = 0 \quad \text{for all } i, j, t. \tag{3.11}$$

The quadratic market model is obtained when factors are the market return r_m and its square r_m^2,

$$r_{i,t} = \beta_{0,i} + \beta_{1,i} r_{m,t} + \beta_{2,i} r_{m,t}^2 + \epsilon_{i,t}. \tag{3.12}$$

The arbitrage pricing theory imposes restrictions on (3.12), predicting the existence of constants λ_1 and λ_2 and following relation among expected returns and betas:

$$\mathbb{E}[r_{i,t}] = \beta_{1,i}\lambda_1 + \beta_{2,i}\lambda_2 \tag{3.13}$$

for any asset i and period t, where λ_k is the expected (excess) return on a portfolio that perfectly mimics the k-th factor, $k = 1, 2$. The traditional way to test whether (3.13) is valid (betas and lambdas are different from zero) is to obtain estimates of $\boldsymbol{\beta}$ for each asset in a time-series regression such as (3.12) and then use those estimates as explanatory variables in a cross-sectional regression of mean returns on betas. To correct for the fact that generated regressors are used in the second step, standard errors must be adjusted, for instance, by the correction proposed by Shanken (1992). A more flexible approach is to estimate betas and lambdas simultaneously in a GMM framework, which avoids the error-in-variables problem of the two-step approach and yields correct (heteroscedastic and autocorrelation consistent) standard errors. From (3.11), (3.12) and (3.13) we set up the following moment conditions:

$$\mathbb{E}[(r_i - \beta_{0,i} - \beta_{1,i} r_m - \beta_{2,i} r_m^2)(1, r_m, r_m^2)'] = 0, \tag{3.14a}$$
$$\mathbb{E}[r_i - \beta_{1,i}\lambda_1 - \beta_{2,i}\lambda_2] = 0, \tag{3.14b}$$

for $i = 1, \ldots, N$. The $3N$ conditions in (3.14a) come from the basic orthogonality assumption (3.11) and exactly identify the beta parameters. The N additional conditions in (3.14b) come from the theoretical pricing model (3.13), allow the estimation of the risk premia and provide $N - 2$ overidentifying restrictions to be tested by Hansen's J statistic.

3.2.3 Stochastic Discount Factor

We can test for a three-moment CAPM without any assumptions about the data generating process of returns, relying only on very general restrictions on preferences. The traditional CAPM implies a pricing kernel $\mathcal{K}$ that is a linear function of the return on aggregate wealth. We can form a second-order[4] Taylor expansion of the SDF that will also depend on squared returns on the market portfolio,

$$\mathcal{K}_{t+1} = \theta_0 + \theta_1 r_{m,t+1} + \theta_2 r_{m,t+1}^2, \tag{3.15}$$

and we expect $\theta_1 < 0$ and $\theta_2 > 0$ as they are proportional to U'' and U''', respectively. However, when excess returns are used, the mean of the SDF is not identified: as $\mathbb{E}_t[\mathcal{K}_{t+1} r_{i,t+1}] = 0$, we have $\mathbb{E}_t[(a\,\mathcal{K}_{t+1}) r_{i,t+1}] = 0$ for any real number a. Following Cochrane (2005), we normalize $\theta_0 = 1$ in order to separately identify θ_1 and θ_2. The fundamental pricing equation becomes $\mathbb{E}_t[(1 + \theta_1 r_{m,t+1} + \theta_2 r_{m,t+1}^2) r_{i,t+1}] = 0$, and if the parameters θ_1 and θ_2 are constant over time, the simplest form of testing this specification of the SDF is by taking the unconditional expectation of this pricing equation and to use the following N moment conditions,

$$\mathbb{E}[(1 + \theta_1 r_m + \theta_2 r_m^2) r_i] = 0. \tag{3.16}$$

to estimate $\boldsymbol{\theta} = (\theta_1 \ \theta_2)'$.

The theoretical models in (3.13) and (3.16) give subtle different answers to the question whether coskewness is a relevant factor for pricing[5]. A statistically significant $\hat{\theta}_k$ in (3.16) indicates that the k-th factor provides useful information about marginal utility beyond the conveyed by the other $(K-1)$ factors, so the price of assets correlated with the k-th factor should account

[4] Dittmar (2002) estimates an SDF that also depends on the cubic return on aggregate wealth. Coupled with a fourth-order expansion of the investor's utility and additional assumptions on his risk-taking behavior (namely, decreasing absolute *prudence*, $-d[U'''/U'']/dW < 0$), the agent exhibits $U'''' < 0$ and aversion to kurtosis in returns.

[5] In fact, the pricing kernel (3.15) and the beta pricing model (3.12)-(3.13) are equivalent representations of the three-moment CAPM. To see this, let $\tilde{\boldsymbol{f}}_t = (\tilde{R}_{m,t} \ \tilde{R}_{m,t}^2)'$ denote demeaned factors, $\boldsymbol{\theta} = (\theta_1 \ \theta_2)'$, $\boldsymbol{\beta}_i = (\beta_{i,1} \ \beta_{i,2})'$ and $\boldsymbol{\lambda} = (\lambda_1 \ \lambda_2)'$. Then, in the case of excess returns, we normalize $\mathcal{K}_{t+1} = 1 + \boldsymbol{\theta}' \tilde{\boldsymbol{f}}_{t+1}$ and have the following relationship among the beta and SDF representations:

$$\boldsymbol{\beta}_i = \mathbb{E}_t[\tilde{\boldsymbol{f}}_{t+1} \tilde{\boldsymbol{f}}'_{t+1}]^{-1} \mathbb{E}_t[\tilde{\boldsymbol{f}}_{t+1} R_{i,t+1}], \tag{3.17a}$$

$$\boldsymbol{\lambda} = -\mathbb{E}_t[\tilde{\boldsymbol{f}}_{t+1} \tilde{\boldsymbol{f}}'_{t+1}] \boldsymbol{\theta} = -\mathbb{E}_t[\mathcal{K}_{t+1} \tilde{\boldsymbol{f}}_{t+1}]. \tag{3.17b}$$

The elements of $\boldsymbol{\lambda}$ are commonly known as factor risk prices, and this can be justified by the term $\mathbb{E}_t[\mathcal{K}_{t+1} \tilde{\boldsymbol{f}}_{t+1}]$ in the right-hand side of (3.17b).

for this risk exposure. On the other hand, a significant $\hat{\lambda}_k$ in (3.13) simply indicates that the k-th factor is correlated with the SDF, but this might be for spurious reasons. Cochrane (2005) interprets θ_k as the coefficient of the k-th factor in a multiple regression of the SDF on all factors, and λ_k as the coefficient of the k-th factor in a simple regression of $\mathcal{K}_{t+1}$ on f_k. The formula for $\boldsymbol{\lambda}$ in (3.17b) can be simplified to $\boldsymbol{\lambda} = \boldsymbol{\Sigma}_f \boldsymbol{\theta}$, so factor prices are linear combinations of θ's, where weights are the covariances among factors. When factors are orthogonal to each other, $\boldsymbol{\Sigma}_f$ is diagonal and $\boldsymbol{\lambda} = \mathbf{0}$ is necessary and sufficient for $\boldsymbol{\theta} = \mathbf{0}$. In most cases, however, factors are correlated so θ and λ have different interpretations. If we let $\tilde{R}_m \equiv R_m - \bar{R}_m$, then $\text{cov}(\tilde{R}_m, \tilde{R}_m^2) = \mathbb{E}[\tilde{R}_m^3]$, so factors are correlated if and only if $\tilde{R}_m$ is skewed. This is an empirical question, and we provide some evidence in Table C.1 in the Appendix. The results show that the returns on our chosen proxy of the market portfolio, the Ibovespa index, had a negatively skewed distribution during the entire 2000-2009 period and during both the first and last halves of that period.

3.3 Empirical Results

Our data comprises monthly and weekly (log) returns of brazilian stocks from 2000:01 to 2009:12, totaling 120 monthly or 520 weekly observations during the last 10 years. This period is characterized by a consistent macroeconomic environment and policy in Brazil, namely of low inflation, floating exchange rates and inflation targeting. We sampled stocks on a yearly basis, and for each year from 2000 to 2009 we selected every company whose most liquid stock was traded in at least 70% of all trading sessions of the year. We formed equally weighted and value weighted portfolios according to size, price-to-book ratio and industryand report all results in tables at the Appendix.

3.3.1 Summary Statistics of Returns

Table C.2 contains summary statistics for the ten equally-weighted size-sorted portfolios of our sample. The sensitivity to squared market returns (coskewness) is significant across all portfolios, even after covariance risk is controlled for. Also, there appears to exist a positive association among size and both factor loadings: rank correlation (Kendall's τ) of $\hat{\beta}_1$ and $\hat{\beta}_2$ with respect to size are 0.51 and 0.56 respectively, and significant at 5%. The positive relation among size and coskewness is consistent with the findings of Harvey and Siddique (2000) and Barone-Adesi, Gagliardini and Urga (2004) in the U.S. market. On the other hand, there is apparently no correlation

among betas and average returns in the cross-section.

In Table C.3 we report summary statistics for the portfolios formed according to price-to-book (PB) ratio deciles. As in the case of size-sorted portfolios, coskewness is highly significant in the cross-section: only stocks in the lowest decile do not appear to have a significant β_2. However, under this ordering criteria the factor loadings β_1 and β_2 do not appear to be correlated with either average returns or price-to-book decile ranks.

Lastly, we show in Table C.4 the summary statistics for industry portfolios. Only four out of the 17 industries did not have a significant exposure to the volatility of the market index. The average realized return is negatively correlated with β_1 across industries in this sample ($\tau = -0.33$ with a p-value of 0.07). This negative relation among size and this correlation disappears when r_m^2 is omitted (the single-index model with r_m is estimated). When squared market returns are included, there is no correlation among β_2 and mean returns in the cross-section.

The general features we learn from the information in Tables C.2, C.3 and C.4 are that the returns on stock portfolios, whether formed by size, price-to-book or industry, are indeed sensitive to market volatility, measured by the squared returns on the benchmark index. Second, the market volatility tends to reduce average *ex-post* returns, since estimates of β_2 are eminently negative. Third, coskewness seems to vary over time, as most portfolios display serial correlation in $\hat{u}_{i,t}\hat{u}^2_{m,t}$, where $\hat{u}_{i,t}$ and $\hat{u}^2_{m,t}$ are the residuals from regressing r_i and r_m^2 on r_m, respectively. The natural question that follows is whether coskewness risk is indeed priced in the market, and consequently whether investors command a premium for holding assets with a significant exposure to this risk factor.

3.3.2 Estimates of the Three-Moment CAPM

The Kraus and Litzenberger / Sears and Wei model represented by (3.7) is the most detailed specification of the three-moment CAPM that we consider. There is a total of $3N + 4$ parameters, where N is the number of assets, and we have a total of $2N - 1$ overidentifying restrictions in our estimation framework. The results for this model are presented in Table C.5 for monthly returns, and Table C.6 for weekly returns. We report estimates for ϕ and S_m under each aggregation (size, price-to-book or industry) and weighting (equally or value-based) criteria, and provide the results of the significance tests of the γ_2 coefficients and of the J test. In spite the fact that the overidentifying moments were not rejected and γ_2 was jointly significant when the model was tested on monthly returns, ϕ was positive and significant for all equally-weighted portfolios, contrary to what the theory predicts.

With weekly data, we failed to find any significant parameter related to the preference for skewness.

The arbitrage pricing model (3.12) and (3.13) is slightly less parametrized than the KL/SW model, and the moment conditions (3.14) provide $N-2$ overidentifying restrictions to be tested. The results are presented in Tables C.7 and C.8. For both monthly and weekly data, we did not find either risk premium to be significative, with some exceptions for a few types of portfolios in certain periods. The coskewness parameters β_2, on the other hand, were almost always jointly significative. Finally, while the pricing errors β_0 were also significative across all portfolios, periods and return horizons, this does not indicate that the our estimated model is, from an APT perspective, misspecified. As noted by Barone-Adesi, Gagliardini and Urga (2004), the squared market return is not a traded portfolio, which implies the intercept restriction $\beta_{0,i} = \beta_{2,i}(\lambda_2 - \mathbb{E}[r_{m,t}^2])$ in the model (3.12).

It is important to notice that in both the KL/SW and the APT models, estimation was not successful in a significant number of cases (each case being a portfolio type and a sample period), as the GMM objective function did not seem to be well-behaved and the numerical optimization procedure[6] failed to find an optimum.

Finally, the less parameterized model is the stochastic discount factor specified in (3.15). There are only two parameters to be estimated, and a total of $N-2$ overidentifying restrictions. We managed to find an optimum for the GMM criterion in every case, and the model passed the J test in 30 out of the 36 cases studied. Despite numerically successful, the evidence was contrary to the prediction of the theory (in particular, $\theta_2 > 0$). In the few cases (5 out of 36, all under weekly returns) where $\hat{\theta}_2$ had the correct theoretical sign, it was not statistically significant.

3.4 Conclusion

In this paper we attempted to perform a comprehensive test of the importance of skewness to price assets in the brazilian market, in the sense that different empirical approaches to the same basic theoretical model were tested on different portfolios in different sample periods, using the GMM as a unifying estimation principle. We found that the return on most stock portfolios are affected by the squared market return, but investors do not seem to require a premium for this risk factor. While the sample size, both in the time-series and in the cross-section dimensions, may not be big enough to allow sufficient power to the statistical tests employed, our approach may

[6]All estimations were conducted with the `gmm` function in Stata 11.1

also be subject to the usual problem in the asset pricing literature of only partially measuring aggregate wealth by considering solely the return on a broad stock portfolio.

Appendices

Appendix A

Table A.1: Average betas of regulated industries (global).

Industry	Number of Firms	Unlevered β
Banks	564	0.24
Power Utilities	596	0.38
Railroads	49	0.39
Water Utilities	83	0.42
Air Transportation	136	0.47
Ground Transportation	193	0.50
Telecom Services	278	0.52
Oil & Gas Distribution	142	0.55

Source: `www.stern.nyu.edu/~adamodar/pc/datasets/betaGlobal.xls`

Table A.2: Historical returns of the U.S. equity market and risk-free rate.

	Accum. returns (1925 = 100)		5-year holding period returns		
Period	T-Bonds	Stocks	r_f	r_m	$r_m - r_f$
1925 - 30	125.76	151.58	25.8%	51.6%	20.5%
1930 - 35	158.76	176.72	26.2%	16.6%	-7.6%
1935 - 40	189.96	181.15	19.7%	2.5%	-14.3%
1940 - 45	208.20	396.45	9.6%	118.9%	99.7%
1945 - 50	222.69	635.94	7.0%	60.4%	50.0%
1950 - 55	239.19	1,855.94	7.4%	191.8%	171.7%
1955 - 60	282.25	2,845.42	18.0%	53.3%	29.9%
1960 - 65	324.17	5,300.43	14.9%	86.3%	62.2%
1965 - 70	415.68	6,246.97	28.2%	17.9%	-8.1%
1970 - 75	566.59	7,314.74	36.3%	17.1%	-14.1%
1975 - 80	725.93	14,052.31	28.1%	92.1%	49.9%
1980 - 85	1,511.59	27,913.82	108.2%	98.6%	-4.6%
1985 - 90	2,362.16	51,754.71	56.3%	85.4%	18.6%
1990 - 95	3,603.14	111,405.99	52.5%	115.3%	41.1%
1995 - 00	4,859.84	258,645.41	34.9%	132.2%	72.1%
2000 - 05	6,268.36	265,782.39	29.0%	2.8%	-20.3%
Average			**31.4%**	71.4%	**34.2%**
Std. dev.			24.9%	**53.6%**	51.1%

Source: Stocks, Bonds, Bills and Inflation 2007 Yearbook. The historical equity premium, $r_m - r_f$, is the compounded return of stocks over the Treasury bonds, $(1+r_m)/(1+r_f)-1$ and not the simple difference.

Table A.3: Feasible parameter domain by the source of restriction.

Parameter	**Model**	**Equilibrium ($\hat{\mathcal{E}}$)**	**Refinement**
d	$d > \mu_\theta + \sigma_\theta$	See (1.15)	$\{7.38, 12, 50\}$
ν	$0 < \nu < 1/\sigma_m$	-	$(0, 0.30]$
μ_θ	$\mu_\theta > \sigma_\theta$	-	Numeraire
σ_θ	$\sigma_\theta > 0$	-	$(0, 0.20]$
ψ	$\psi > 1 + \lambda$	See (1.16) and (1.19)	$\{\psi_L(d), \tilde{\psi}(d), \psi_H(d)\}$
ρ	$-1 \leq \rho \leq 1$	-	$[-0.50, 0.50]$

A.1 Regulation under Complete Information

A.1.1 Optimal Payment Profile

Consider the choice of a policy (1.5) that maximizes (1.8) subject to (1.3), (1.4) and (1.6) but taking I as exogenous or pre-determined. In this problem there is a complete separation between the choice of output and effort from the design of the payment mechanism: the optimal (unconstrained) production and effort levels maximize expected net consumer surplus

$$\sum_{j=L,H}\sum_{k=L,H} P(\theta_j, m_k)\{S(q_{jk}) - (1+\lambda)[(\theta_j - e_{jk})q_{jk} + \Psi(e_{jk})]\},$$

and are easily obtained from the first-order conditions. The optimal payment profile minimizes the expected rent

$$\sum_j\sum_k P(\theta_j, m_k)U_{jk} \propto (1+\rho)(U_{LL}+U_{HH}) + (1-\rho)(U_{LH}+U_{HL})$$

subject to the valuation restriction (1.6), which can be written as

$$\begin{aligned}(1+\rho)[U_{LL}(1+\gamma/\sigma_m) + U_{HH}(1-\gamma/\sigma_m)]+\\ (1-\rho)[U_{HL}(1+\gamma/\sigma_m) + U_{LH}(1-\gamma/\sigma_m)] \geq 4I(1+r_f).\end{aligned}$$

First, it is clear that (1.6) must be active at any solution, or rent could be decreased in at least some state, reducing the expected rent. Also, inspection of (1.6) shows that it is never optimal to leave rent in good states, when $k = H$, as transferring cash flow from U_{jH} to $U_{j'L}$, for $j, j' = L, H$ and $j \neq j'$ keeps expected rent unchanged and slacks the restriction. Thus, $U_{LH} = U_{HH} = 0$ at any solution, which effectively reduces the problem to the minimization of $(1+\rho)U_{LL} + (1-\rho)U_{HL}$ subject to

$$(1+\rho)U_{LL} + (1-\rho)U_{HL} = \frac{4I(1+r_f)}{1+\gamma/\sigma_m},$$

which leads to (1.12).

A.1.2 Optimal Investment

In Appendix A.1.1 we characterized the optimal payment mechanism for any exogenous level of investment I. In this section we derive the optimal determination of I under the payment scheme (1.12). Without loss of generality, we can set $U^*_{HL} = 0$ and concentrate all rent in the (θ_L, m_L) state.

In the unconstrained optimum (1.9), the maximum output takes place in the (θ_L, m_H) state. Consider the case in which only the capacity restriction in this state is active. Then, the (restricted) optimal output profile is the same as (1.9b) for all other states and it is equal to

$$I^* = q^*_{LH} - 4\lambda \frac{\psi}{\psi - (1+\lambda)} \frac{\sigma_m}{\gamma + \sigma_m} \frac{1 + r_f}{1 - \rho}$$

when cost is low and demand is high. For this production plan to be feasible, and if $d\,\nu\,\sigma_m \geq (1+\lambda)\sigma_\theta$, output in state (θ_L, m_L) must be lower than I^*, which requires $2\lambda(1+r_f) \leq d\nu(\gamma + \sigma_m)(1-\rho)$. On the other hand, if $d\,\nu\,\sigma_m < (1+\lambda)\sigma_\theta$, then output in state (θ_H, m_H) must be lower than I^*, which requires $2\lambda(1+r_f)\sigma_m \leq (1+\lambda)(\gamma+\sigma_m)(1-\rho)\sigma_\theta$. If these conditions are met, then I^* is the optimal capacity level and output in the (θ_L, m_H) state. Otherwise, the analysis is more complex and depend on different combinations of the exogenous parameters.

A.2 Regulation under Asymmetric Information

Under asymmetric information, the truth-telling conditions (1.17) are obtained as follows. Let $E(C, \theta, q)$ be the effort required for a type-θ firm to produce q at cost C, or equivalently, $C = C(q, E(C, \theta, q), \theta)$. When a θ_j−firm announces its type as being $\theta_{j'}$, her payoff is $U(\theta_j, \theta_{j'}) = t(\theta_j) - \Psi(\theta_j - c_{j'k})$, where $c_{jk} \equiv C_{jk}/q_{jk}$ is the marginal cost and $j, j' = L, H$. Then, a menu $\{(t_{jk}, C_{jk}, q_{jk})_{j,k=L,H}\}$ induces truth-telling for the efficient firm if and only if

$$t_{Lk} - \Psi(\theta_L - c_{Lk}) \geq t_{Hk} - \Psi(\theta_L - c_{Hk}).$$

Since $\theta_L = \theta_H - 2\sigma_\theta$, the above condition can be restated as $U_{Lk} \geq t_{Hk} - \Psi(\theta_H - c_{Hk} - 2\sigma_\theta)$. As $\theta_H - c_{Hk} = e_{Hk}$, adding and subtracting $\Psi(e_{Hk})$ in the right-hand side yields $U_{Lk} \geq U_{Hk} + \Psi(e_{Hk}) - \Psi(e_{Hk} - 2\sigma_\theta)$ as desired. The incentive constraints for the inefficient firm are obtained analogously.

We wish to maximize (1.8) subject to (1.3), (1.4), (1.6) and (1.17). First, notice that incentive compatibility requires $U_{LH} \geq U_{HH}$. Second, we must have $U_{HH} = 0$ at any solution. Otherwise, it is feasible to reduce U_{HH} and U_{LH} each by an ε such that $U_{HH} \geq \varepsilon > 0$ and to increase U_{LL} and U_{HL} by $\varepsilon(1-\gamma/\sigma_m)/(1+\gamma/\sigma_m) < \varepsilon$ each, which keeps the financial and incentive restrictions unchanged and strictly decreases $\mathbb{E}[U]$, increasing expected welfare. Then, U_{LL}, U_{LH} and U_{HL} solve the system of equations formed by (1.17a), (1.17c) and (1.6), and are substituted back into the objective function, which is then maximized for the optimal levels of effort and output. The ignored

restrictions are checked *ex-post* whether are satisfied at the candidate solution. A necessary and sufficient condition for incentive compatibility for the inefficient firm in state m_k is $e_{Lk} + 2\sigma_\theta \geq e_{Hk}$, and this is indeed the case for all k under the solution in Proposition 1.

A.2.1 Cost Unobservability

The regulator designs a transfer-output menu $\{(t_{jk}, q_{jk})_{j=L,H}\}$ for each $k = L, H$. A firm with cost θ_j that selects the contract $(t_{j'k}, q_{j'k})$ enjoys rent

$$U_{jk}(q_{j'k}, E(q_{j'k})) = t_{j'k} - C(q_{j'k}, E(q_{j'k}), \theta_j) - \Psi(E(q_{j'k}))$$

for $j, j' = L, H$. Incentive compatibility requires

$$U_{jk}(q_{jk}, E(q_{jk})) \geq U_{jk}(q_{j'k}, E(q_{j'k}))$$

for all $j, j' = L, H$ and $j \neq j'$. If we let $U_{jk} \equiv U_{jk}(q_{jk}, E(q_{jk}))$, then adding and subtracting $C(q_{j'k}, E(q_{j'k}), \theta_{j'})$ in the right side of the IC condition above results in

$$U_{jk} \geq U_{j'k} + q_{j'k}(\theta_{j'} - \theta_j)$$

for all $j, j' = L, H$ and $j \neq j'$, resulting in the constraints (1.21). By a similar argument as in the incomplete information case with cost observability, $U_{HH} = 0$ at any solution. U_{LL}, U_{LH} and U_{HL} are found by solving the system formed by (1.21a), (1.21b) and (1.6) and substituted back into the objective function.

A.2.2 Price-Cap Regulation

From (1.22) and $U_{jk}(p) \equiv U_{jk}(p, E_k(p))$, we have

$$U_{jk}(p) = d(1 + \nu m_k)[d(1 + \nu m_k) - 2\psi\theta_j + p(2\psi - 1)](2\psi)^{-1}.$$

Some algebra shows that

$$U_{Lk}(p) - U_{Hk}(p) = 2\sigma_\theta[d(1 + \nu m_k) - p], \tag{A.1a}$$
$$U_{jH}(p) - U_{jL}(p) = 2d\nu\sigma_m\psi^{-1}[d + p(\psi - 1) - \theta_j\psi]. \tag{A.1b}$$

Then, (A.1a) is positive for all k as any equilibrium price-cap satisfies $p < d(1 + \nu m_k)$ otherwise demand would be zero. And (A.1b) is positive for all j as long as

$$p \geq \frac{\theta_j\psi - d}{\psi - 1}. \tag{A.2}$$

The lower bound (A.2) for p increases with ψ and θ, and decreases with d. So evaluating (A.2) at $\theta_j = \theta_H$, d as the infimum of (1.15) and taking the limit as $\psi \to \infty$, we have

$$\lim_{\psi \to \infty} \frac{\theta_H \psi - [(1+\lambda)\theta_H/(1-\nu\sigma_m)]}{\psi - 1} = \theta_H,$$

which means that the highest possible lower bound for p is θ_H. This must be trivially satisfied by any equilibrium price cap, otherwise the firm is likely to be valued at less than the necessary investment and no production would take place. We will assume that in equilibrium $p > \theta_H$ and check whether this condition holds *ex-post*. Then, it is trivial to show that if (A.1b) and (A.1a) are positive, $U_{HL}(p) \leq U_{jk}(p)$ for all j, k as $U_{HL}(p) < U_{LL}(p)$ by (A.1a), $U_{HL}(p) < U_{HH}(p)$ by (A.1b) and $U_{HH}(p) < U_{LH}(p)$ by (A.1a).

Appendix B

B.1 Moments of the Return Process

When y_t is written as a function of current and past shocks, $y_t = f(\epsilon_t, \epsilon_{t-1}, \ldots)$, the representation in (2.5) can be formalized as a first-order Taylor expansion of $f(\cdot)$ around $\epsilon_t = 0$ given past shocks $\{\epsilon_{t-1}, \epsilon_{t-2}, \ldots\}$, and yields

$$y_t = f(0, \epsilon_{t-1}, \epsilon_{t-2}, \ldots) + f_1(0, \epsilon_{t-1}, \epsilon_{t-2}, \ldots)\epsilon_t, \tag{B.1}$$

where $f_1(\cdot)$ is the derivative of f with regards to ϵ_t. From (B.1) it is easy to derive the conditional moments of y_t. The conditional mean is given by

$$\mu_t \equiv \mathbb{E}_{t-1} y_t = f(0, \epsilon_{t-1}, \epsilon_{t-2}, \ldots),$$

and the conditional variance is

$$h_t^2 \equiv \mathbb{E}_{t-1}(y_t - \mathbb{E}_{t-1} y_t)^2 = f_1(\cdot)^2 \mathbb{E}_{t-1}\epsilon_t^2 = f_1(\cdot)^2.$$

The specification in (B.1) also implies that all higher-order conditional moments of y_t are linked to the function $f_1(\cdot)$ and to the error process $\{\epsilon_t\}$. For all $k \geq 2$,

$$\mathbb{E}_{t-1}(y_t - \mathbb{E}_{t-1} y_t)^k = f_1(\cdot)^k \mathbb{E}_{t-1}\epsilon_t^k = h_t^k \mathbb{E}\epsilon_t^k,$$

where in the last equality the conditional expectation was substituted by the unconditional one as ϵ_t is a i.i.d. process by assumption.We could go one step further and form a second-order expansion of $f(\cdot)$ around $\epsilon_t = 0$ by adding the term $(1/2)f_{11}(0, \epsilon_{t-1}, \epsilon_{t-2}, \ldots)\epsilon_t^2$ to (B.1), where $f_{11}(\cdot)$ is the second derivative of f w.r.t. ϵ_t. The potential gain in flexibility however has a high price in terms of tractability, as the conditional variance and all higher-order moments of $\{y_t\}$ will depend on cross-products of powers of two unknown functions, $h(\cdot)$ and $f_{11}(\cdot)$, and such an extension does not seem very

promising[1].

The above first-order approximation implies that the conditional skewness of $\{y_t\}$ depends solely on the *un*conditional third moment of the disturbance,

$$\frac{\mathbb{E}_{t-1}(y_t - \mathbb{E}_{t-1}[y_t])^3}{[\mathbb{E}_{t-1}(y_t - \mathbb{E}_{t-1}[y_t])^2]^{3/2}} = \frac{h_t^3 \mathbb{E}_{t-1}\epsilon_t^3}{[h_t^2 \mathbb{E}_{t-1}\epsilon_t^2]^{3/2}} = \mathbb{E}\epsilon_t^3. \tag{B.2}$$

In (B.2) it becomes clear that symmetry of ϵ_t is not an innocuous assumption, as it implies conditional symmetry of the return series y_t.

The unconditional moments of y_t in model in (2.5) are not as simple as the conditional ones. The unconditional mean is

$$\mathbb{E}y_t = \mathbb{E}\mu_t,$$

the unconditional variance is

$$\begin{aligned}
\mathbb{E}[(y_t - \mathbb{E}y_t)^2] &= \mathbb{E}[y_t^2 - 2y_t\mathbb{E}\mu_t + (\mathbb{E}\mu_t)^2] \\
&= \mathbb{E}[(\mu_t + \eta_t)^2 - 2(\mu_t + \eta_t)\mathbb{E}\mu_t + (\mathbb{E}\mu_t)^2] \\
&= \mathbb{E}[\mu_t^2 + \eta_t^2 - 2\mu_t\mathbb{E}\mu_t + (\mathbb{E}\mu_t)^2] \\
&= \mathbb{E}\mu_t^2 - (\mathbb{E}\mu_t)^2 + \mathbb{E}\eta_t^2 \\
&= \operatorname{var}\mu_t + \mathbb{E}\eta_t^2 = \operatorname{var}\mu_t + \mathbb{E}h_t^2 \qquad \blacksquare
\end{aligned}$$

and the unconditional third moment is

$$\begin{aligned}
\mathbb{E}[(y_t - \mathbb{E}y_t)^3] &= \mathbb{E}[y_t^3 - 3y_t^2\mathbb{E}\mu_t + 3y_t(\mathbb{E}\mu_t)^2 - (\mathbb{E}\mu_t)^3] \\
&= \mathbb{E}[(\mu_t + \eta_t)^3 - 3(\mu_t + \eta_t)^2\mathbb{E}\mu_t + 3(\mu_t + \eta_t)(\mathbb{E}\mu_t)^2 - (\mathbb{E}\mu_t)^3] \\
&= \mathbb{E}[\mu_t^3 - 3\mu_t^2\mathbb{E}\mu_t + 3\mu_t(\mathbb{E}\mu_t)^2 - (\mathbb{E}\mu_t)^3] + 3\mathbb{E}[\mu_t\eta_t^2 - \mathbb{E}\mu_t\eta_t^2] \\
&+ \mathbb{E}[3(\mathbb{E}\mu_t)^2\eta_t + 3\mu_t^2\eta_t - 6\mu_t\eta_t\mathbb{E}\mu_t] + \mathbb{E}\eta_t^3 \\
&= \mathbb{E}[(\mu_t - \mathbb{E}\mu_t)^3] + 3\operatorname{cov}(\mu_t, \eta_t^2) + \mathbb{E}\eta_t^3 \qquad \blacksquare
\end{aligned}$$

where $\mathbb{E}[3(\mathbb{E}\mu_t)^2\eta_t + 3\mu_t^2\eta_t - 6\mu_t\eta_t\mathbb{E}\mu_t] = 0$ since $\eta_t = h_t\epsilon_t$ and ϵ_t is an i.i.d. shock with unit variance. The unconditional skewness of y_t comes from three sources: the conditional mean, which may be skewed by itself or linearly related to the squared disturbance; the conditional standard deviation, which may or may not respond asymmetrically to shocks; and from disturbance itself, which may be asymmetrically distributed. For this reason, to investigate asymmetry in the distribution of returns, one must be careful

[1]Identification in such a model would quickly become a nightmare. For instance, the conditional variance would be given by $f_1(\cdot)^2 + 2f_1(\cdot)f_{11}(\cdot) + f_{11}(\cdot)^2\mathbb{E}_{t-1}\epsilon_t^4$, and the conditional third moment would be $f_1(\cdot)^3\mathbb{E}_{t-1}\epsilon_t^3 + 3f_1(\cdot)^2 f_{11}(\cdot)\mathbb{E}_{t-1}\epsilon_t^4 + 3f_1(\cdot)f_{11}(\cdot)^2\mathbb{E}_{t-1}\epsilon_t^5 + f_{11}(\cdot)^3\mathbb{E}_{t-1}\epsilon_t^6$.

when imposing restrictions on the three major ingredients of the model (the conditional mean and variance and the error process), since they have a direct influence on the hypothesis being tested. In particular, the variance equation and the disturbance process should be as general as possible, which is the reason we choose the specifications in (2.12) and (2.13).

B.2 The Skewed-t Distribution

Fernandez and Steel (1998) show that for any density $f(\cdot)$ continuous, symmetric at zero and unimodal, the transformation

$$g(z\,|\,\xi) = \frac{2}{\xi + 1/\xi}\left[f(z/\xi)\mathbb{I}_{\{z\geq 0\}} + f(z\,\xi)\mathbb{I}_{\{z<0\}}\right] \tag{B.3}$$

for any $\xi \in (0,\infty)$ is a density based on $f(\cdot)$ that retains unimodality at zero for all ξ but loses symmetry whenever $\xi \neq 1$. We take f to be the standardized Student-t density,

$$f(z_t) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{(\nu-2)\pi}}\left(1+\frac{z_t^2}{\nu-2}\right)^{-(\nu+1)/2}. \tag{B.4}$$

The symmetry parameter ξ is proportional to the relative mass of positive values, as $P(x \geq 0\,|\,\xi)/P(x < 0\,|\,\xi) = \xi^2$. The sign of $\log(\xi)$ indicates the sign (or direction) of the asymmetry in $g(\cdot)$: when $\xi \in (0,1)$, $\log(\xi) < 0$ and $g(\cdot)$ is negatively (of left) skewed; and when $\xi \in (1,\infty)$, $\log(\xi) > 0$ and $g(\cdot)$ is positively (of right) skewed. Figure B.1 illustrates the role of ξ in controlling the skewness of $g(\cdot)$.

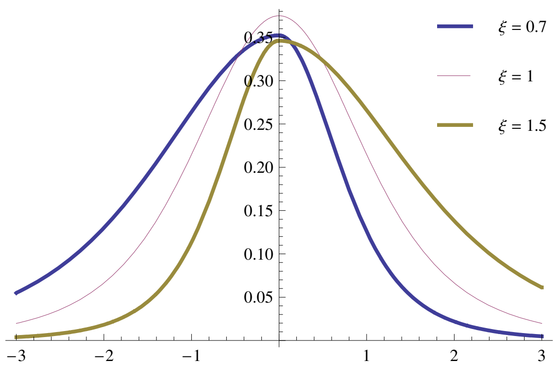


Figure B.1: Skewed versions of a Student-t Density

The (uncentered) moments of z under $g(\cdot)$ are given by

$$\mathbb{E}[z^r \mid \xi] = \int_0^\infty s^r 2f(s)ds \, \frac{\xi^{r+1} + (-1)^r \xi^{-(r+1)}}{\xi + 1/\xi}. \tag{B.5}$$

The skewness coefficient (2.6) is a complicated function under the previous formula, but the asymmetry measure introduced by Arnold and Groenenveld (1995), one minus two times the probability mass left of the mode, can be calculated as

$$AG(x \mid \xi) = \frac{\xi^2 - 1}{\xi^2 + 1}, \tag{B.6}$$

and displays several desirable properties: it is an increasing function of ξ; preserves the convex orderings of distributions of van Zwet (1964) when $f(\cdot)$ is differentiable; takes values on $(-1, 1)$ and vanishes when $\xi = 1$. Therefore, (B.6) is a bounded measure of asymmetry, with a very intuitive interpretation.

B.3 Summary of Results

Table B.1: Firms with significant skewness by the (2.9) and (2.11) tests.

			Skewness $\widehat{S}$		Robust $S_R(0.25)$	
Period	Type of Firm	Total	Positive	Negative	Positive	Negative
1994-1997	Regulated	28	22	0	12	0
	Unregulated	59	45	2	10	10
1997-2000	Regulated	29	10	13	4	1
	Unregulated	50	35	5	9	8
2000-2003	Regulated	37	13	5	1	4
	Unregulated	50	30	5	11	3
2003-2006	Regulated	42	29	1	11	2
	Unregulated	54	32	1	21	1
2006-2009	Regulated	47	22	6	6	4
	Unregulated	85	35	21	11	8
	Total	481	273	59	96	41

Note: Entries are the number of companies with statistically significant $\widehat{S}$ and $\widehat{S}_R(0.25)$, assuming i.i.d. normal returns and a 10% significance level.

Table B.2: Results of the asymmetry tests described in Section 2.2.1, at 5% nominal size

	Sampled Firms		Unconditional Tests								
Period	Type	Quantity	$\widehat{S}(+)$	$\widehat{S}(-)$	HS	HC	HP	KS	KC	KP	$\widehat{S}_\rho$
1994-1997	Regulated	28	8	0	3	13	8	15	15	15	12
	Unregulated	59	26	0	17	30	30	28	30	33	23
1997-2000	Regulated	29	3	2	1	2	2	2	1	1	8
	Unregulated	50	18	0	10	16	15	5	5	9	20
2000-2003	Regulated	37	9	0	6	7	7	10	8	9	7
	Unregulated	50	15	1	8	18	16	15	16	15	13
2003-2006	Regulated	42	16	0	6	17	18	11	16	16	12
	Unregulated	54	22	0	13	22	22	17	18	17	19
2006-2009	Regulated	47	3	0	2	6	4	3	5	4	5
	Unregulated	85	14	2	8	16	16	8	8	8	11
Total		481	134	4	74	147	138	114	122	127	130

Note: Entries are the number of firms with asymmetric returns according to the tests proposed by Lisi (the skewness coefficient $\widehat{S}$ with bootstrapped standard errors), Chen and Lin (the H and K tests), and Maasoumi and Racine (the $\widehat{S}_\rho$ test), all at a 5% level of significance. In this table only the $\widehat{S}$ test is capable of distinguishing positive (right-skewed) from negative (left-skewed) asymmetry. Its results are reported with indicative (+) and (−) signs.

Table B.3: Results of the asymmetry tests described in Section 2.2.1, at 10% nominal size

	Sampled Firms		Unconditional Tests								
Period	Type	Quantity	$\widehat{S}(+)$	$\widehat{S}(-)$	HS	HC	HP	KS	KC	KP	$\widehat{S}_\rho$
1994-1997	Regulated	28	12	0	4	17	14	19	19	18	19
	Unregulated	59	32	0	24	40	36	33	38	40	25
1997-2000	Regulated	29	3	3	1	6	2	4	3	3	12
	Unregulated	50	23	0	16	25	19	13	16	14	22
2000-2003	Regulated	37	11	0	8	8	9	10	10	10	12
	Unregulated	50	23	1	13	22	22	15	18	20	16
2003-2006	Regulated	42	27	0	9	23	22	16	21	19	13
	Unregulated	54	29	0	19	27	28	22	21	20	21
2006-2009	Regulated	47	10	0	2	7	8	3	10	8	8
	Unregulated	85	22	3	14	21	20	13	11	10	18
Total		481	192	7	110	196	180	148	167	162	166

Note: Entries are the number of firms with asymmetric returns according to the tests proposed by Lisi (the skewness coefficient $\widehat{S}$ with bootstrapped standard errors), Chen and Lin (the H and K tests), and Maasoumi and Racine (the $\widehat{S}_\rho$ test), all at a 10% level of significance. In this table only the $\widehat{S}$ test is capable of distinguishing positive (right-skewed) from negative (left-skewed) asymmetry. Its results are reported with indicative (+) and (−) signs.

Table B.4: Results of the asymmetry tests described in Section 2.2.2, at 5% and 10% nominal sizes

	Sampled Firms		Conditional Tests at 5%				Conditional Tests at 10%			
Period	Type	Quantity	$\widehat{\xi}_i(+)$	$\widehat{\xi}_i(-)$	CS	$\widehat{S}_\rho$	$\widehat{\xi}_i(+)$	$\widehat{\xi}_i(-)$	CS	$\widehat{S}_\rho$
1994-1997	Regulated	28	12	0	19	17	17	0	21	17
	Unregulated	59	16	1	40	21	23	1	43	26
1997-2000	Regulated	29	2	0	10	6	6	0	13	10
	Unregulated	50	12	0	35	15	15	1	38	21
2000-2003	Regulated	37	4	1	10	4	5	1	12	8
	Unregulated	50	8	0	26	12	11	0	29	19
2003-2006	Regulated	42	16	0	15	12	22	0	21	24
	Unregulated	54	17	0	28	23	22	0	32	26
2006-2009	Regulated	47	7	2	12	7	10	3	15	9
	Unregulated	85	12	0	36	22	24	2	43	29
Total		481	106	4	231	139	155	8	267	189

Note: Entries are the number of firms with asymmetric returns according to the tests proposed by Lisi (the skewness coefficient $\widehat{S}$ with bootstrapped standard errors), Bai and Ng (the conditional symmetry CS test) and the QMLE estimate of the asymmetry parameter of the GARCH residual $\widehat{\xi}_i$, at confidence levels 5% and 10%. In this table only the $\widehat{\xi}_i$ test is capable of distinguishing positive (right-skewed) from negative (left-skewed) asymmetry, and their results are reported with indicative (+) and (−) signs.

Table B.5: True contingency table under observability of S

	Asymmetric	Symmetric	Total
Regulated	x	$m-x$	n_1
Unregulated	$k-x$	$n-(k-x)$	n_0
Total	k	$m+n-k$	n_0+n_1

Table B.6: Observed tabulation based on an estimator $\widehat{S}$ of S

	Asymmetric	Symmetric
Reg.	$A \equiv x(1-\beta)+(n_1-x)\alpha$	$B \equiv (n_1-x)(1-\alpha)+x\beta$
Unreg.	$C \equiv (k-x)(1-\beta)+(n_0-k+x)\alpha$	$D \equiv (n_0-k+x)(1-\alpha)+k\beta$
Total	$k(1-\beta)+(n_0+n_1-k)\alpha$	$(n_0+n_1-k)(1-\alpha)+k\beta$

Table B.7: Odds ratio by period and asymmetry test performed. Power = 95%

Test	Period (5% significance level)					Period (10% significance level)				
	94-97	97-00	00-03	03-06	06-09	94-97	97-00	00-03	03-06	06-09
$\widehat{S}$	0.43	$0.31^{\bullet}$	0.71	0.81	$0.12^{\bullet}$	0.59	$0.22^{\bullet}$	$0.41^{\bullet}$	1.66	0.48
HS	$0.23^{\bullet}$	NA	1.14	0.37	NA	$0.06^{\bullet}$	NA	0.71	$0.40^{\bullet}$	NA
HC	0.84	$0.09^{\bullet}$	$0.38^{\bullet}$	1.07	0.52	0.74	$0.19^{\bullet}$	$0.24^{\bullet}$	1.18	0.35
HP	$0.33^{\bullet}$	$0.10^{\bullet}$	0.46	1.18	$0.25^{\bullet}$	0.56	NA	$0.29^{\bullet}$	1.00	0.52
KS	1.27	0.86	0.83	0.74	0.44	1.77	$0.15^{\bullet}$	0.74	0.92	NA
KC	1.11	NA	0.55	1.21	1.38	1.16	$0.00^{\bullet}$	0.50	1.81	$3.95^{\bullet\star}$
KP	0.91	NA	0.71	1.32	0.90	0.86	$0.00^{\bullet}$	$0.42^{\bullet}$	1.36	3.82
$\widehat{S}_{\rho}$	1.26	0.48	0.69	0.71	0.76	$3.26^{\bullet\star}$	0.92	1.05	0.58	0.63
$\widehat{\xi}_i$	$2.18^{\star}$	$0.15^{\bullet}$	0.89	1.32	1.47	$2.76^{\bullet\star}$	0.46	0.55	1.83	0.88
CS	1.10	$0.21^{\bullet}$	$0.30^{\bullet}$	$0.47^{\bullet}$	$0.44^{\bullet}$	1.02	$0.20^{\bullet}$	$0.30^{\bullet}$	0.68	$0.37^{\bullet}$
$\widehat{S}_{\rho}$	$2.97^{\bullet\star}$	0.54	$0.21^{\bullet}$	0.48	0.39	$2.39^{\bullet\star}$	0.63	$0.31^{\bullet}$	1.40	$0.31^{\bullet}$

Notes: Entries are the odds ratio (2.20) associated with a contingency table (see Table B.5). This table shows results of each asymmetry test considered in sections 2.2.1 and 2.2.2, adjusted for the effect (i.e., size and power) of the statistical test, as shown in Table B.6. As the level and power of each test is not known, we tested whether the odds ratio was greater than unity for two different rejection probabilities under the null (5% and 10%). Marks that indicate significance are: $\bullet$ = two-sided alternative ($OR \neq 1$), $\star$ =one-sided alternative ($OR > 1$).

Table B.8: Odds ratio by period and asymmetry test performed. Power = 80%

Test	Period (5% significance level)					Period (10% significance level)				
	94-97	97-00	00-03	03-06	06-09	94-97	97-00	00-03	03-06	06-09
$\hat{S}$	0.43	0.32$^{\bullet}$	0.79	0.96	0.09$^{\bullet}$	0.52	0.15$^{\bullet}$	0.30$^{\bullet}$	2.31$^{\bullet\star}$	0.56
HS	0.16$^{\bullet}$	NA	1.19	0.39	NA	0.10$^{\bullet}$	NA	0.62	0.37$^{\bullet}$	NA
HC	0.69	0.07$^{\bullet}$	0.33$^{\bullet}$	0.98	0.52	0.51	0.12$^{\bullet}$	0.21$^{\bullet}$	1.33	0.26$^{\bullet}$
HP	0.31$^{\bullet}$	0.08$^{\bullet}$	0.42$^{\bullet}$	1.08	0.19$^{\bullet}$	0.50	NA	0.30$^{\bullet}$	0.94	0.52
KS	1.32	0.56	0.82	0.74	0.35	2.34*	0.27	0.83	0.85	NA
KC	1.15	NA	0.49	1.40	1.48	1.30	0.00$^{\bullet}$	0.53	1.93*	5.53$^{\bullet\star}$
KP	0.86	NA	0.79	1.51	0.71	0.84	0.00$^{\bullet}$	0.45	1.56	4.88$^{\bullet\star}$
$\hat{S}_\rho$	1.18	0.53	0.60	0.65	0.79	5.35$^{\bullet\star}$	0.88	1.02	0.65	0.66
$\hat{\xi}_i$	2.09*	0.12$^{\bullet}$	0.64	1.51	1.77	3.13$^{\bullet\star}$	0.34	0.41	1.83	0.82
CS	0.83	0.09$^{\bullet}$	0.26$^{\bullet}$	0.40$^{\bullet}$	0.39$^{\bullet}$	1.47	0.04$^{\bullet}$	0.23$^{\bullet}$	0.56	0.35$^{\bullet}$
$\hat{S}_\rho$	4.30$^{\bullet\star}$	0.51	0.25$^{\bullet}$	0.45$^{\bullet}$	0.37$^{\bullet}$	2.56$^{\bullet\star}$	0.62	0.29$^{\bullet}$	1.59	0.29$^{\bullet}$

Notes: Entries are the odds ratio (2.20) associated with a contingency table (see Table B.5). This table shows results of each asymmetry test considered in sections 2.2.1 and 2.2.2, adjusted for the effect (i.e., size and power) of the statistical test, as shown in Table B.6. As the level and power of each test is not known, we tested whether the odds ratio was greater than unity for two different rejection probabilities under the null (5% and 10%). Marks that indicate significance are: $\bullet$ = two-sided alternative ($OR \neq 1$), $\star$ =one-sided alternative ($OR > 1$).

Table B.9: Odds ratio by period and asymmetry test performed. Power = 65%

Test	Period (5% significance level)					Period (10% significance level)				
	94-97	97-00	00-03	03-06	06-09	94-97	97-00	00-03	03-06	06-09
$\hat{S}$	$0.36^{\bullet}$	$0.25^{\bullet}$	0.67	0.83	$0.07^{\bullet}$	$0.40^{\bullet}$	$0.15^{\bullet}$	$0.24^{\bullet}$	$10.30^{\bullet\star}$	0.50
HS	$0.19^{\bullet}$	NA	1.06	$0.34^{\bullet}$	NA	$0.06^{\bullet}$	NA	0.71	$0.34^{\bullet}$	NA
HC	0.66	$0.05^{\bullet}$	$0.33^{\bullet}$	1.01	0.48	NA	$0.11^{\bullet}$	$0.17^{\bullet}$	1.63	$0.25^{\bullet}$
HP	$0.21^{\bullet}$	$0.05^{\bullet}$	$0.41^{\bullet}$	1.03	$0.24^{\bullet}$	$0.19^{\bullet}$	NA	$0.23^{\bullet}$	1.01	0.45
KS	2.00	0.41	0.91	0.70	0.29	NA	$0.18^{\bullet}$	0.69	0.74	NA
KC	1.43	NA	$0.44^{\bullet}$	1.30	1.04	NA	$0.00^{\bullet}$	0.46	$2.59^{\bullet\star}$	$5.40^{\bullet\star}$
KP	0.83	NA	0.67	1.51	0.9	NA	$0.00^{\bullet}$	$0.34^{\bullet}$	$1.93^{\star}$	$3.95^{\bullet\star}$
$\hat{S}_{\rho}$	1.41	0.45	0.63	0.57	0.57	NA	0.95	1.02	0.57	0.59
$\hat{\xi}_i$	$2.78^{\bullet\star}$	$0.08^{\bullet}$	0.71	1.51	1.69	$10.01^{\bullet\star}$	$0.40^{\bullet}$	0.43	$2.35^{\bullet\star}$	0.78
CS	NA	NA	$0.18^{\bullet}$	$0.32^{\bullet}$	$0.31^{\bullet}$	NA	NA	$0.10^{\bullet}$	$0.36^{\bullet}$	$0.25^{\bullet}$
$\hat{S}_{\rho}$	$12.25^{\bullet\star}$	0.58	$0.29^{\bullet}$	$0.37^{\bullet}$	$0.40^{\bullet}$	$8.14^{\bullet\star}$	0.59	$0.28^{\bullet}$	$2.73^{\bullet\star}$	$0.27^{\bullet}$

Notes: Entries are the odds ratio (2.20) associated with a contingency table (see Table B.5). This table shows results of each asymmetry test considered in sections 2.2.1 and 2.2.2, adjusted for the effect (i.e., size and power) of the statistical test, as shown in Table B.6. As the level and power of each test is not known, we tested whether the odds ratio was greater than unity for two different rejection probabilities under the null (5% and 10%). Marks that indicate significance are: $\bullet$ = two-sided alternative ($OR \neq 1$), $\star$ =one-sided alternative ($OR > 1$).

Table B.10: Proportion of tables with odds ratio different from unity

	Alternative: $OR \neq 1$			Alternative: $OR > 1$		
Panel A: Unconditional Asymmetry						
	$\beta = 5\%$	$\beta = 20\%$	$\beta = 35\%$	$\beta = 5\%$	$\beta = 20\%$	$\beta = 35\%$
$\alpha = 5\%$	22%	25%	33%	0%	0%	0%
$\alpha = 10\%$	36%	39%	58%	6%	17%	16%
Panel B: Conditional Asymmetry						
	Alternative: $OR \neq 1$			Alternative: $OR > 1$		
	$\beta = 5\%$	$\beta = 20\%$	$\beta = 35\%$	$\beta = 5\%$	$\beta = 20\%$	$\beta = 35\%$
$\alpha = 5\%$	47%	60%	69%	13%	13%	15%
$\alpha = 10\%$	47%	47%	77%	13%	13%	31%

Note: Entries are the proportion of odds ratios from Tables B.7 to B.9 that are statistically different from unity, for each pair (α, β).

Table B.11: Randomness in the sequence of $\{r_{(i)}\}$ ordered by asymmetry

Test	**Period** 1994-1997	1997-2000	2000-2003	2003-2006	2006-2009
Panel A: Unconditional Tests					
$\widehat{S}$	0.23	0.53	**0.01**	0.40	0.85
HS	**0.02**	0.25	0.89	0.25	0.95
HC	**0.07**	**0.05**	**0.02**	0.56	0.92
HP	0.93	**0.08**	0.29	0.40	0.68
KS	0.89	0.25	**0.05**	0.25	0.54
KC	0.84	**0.00**	0.78	0.92	0.54
KP	0.40	**0.02**	0.54	0.25	0.80
$\widehat{S}_\rho$	0.16	**0.08**	0.29	0.72	0.54
Panel B: Conditional Tests					
$\widehat{\xi}_i$	0.31	0.79	0.11	**0.01**	0.25
CS	0.84	0.13	0.37	0.64	0.46
$\widehat{S}_\rho$	0.40	**0.01**	0.45	0.32	0.46

Note: Entries are the one-sided p-values of the runs test on $\{r_{(i)}\}$ for each asymmetry measure and sub-period, where p-values lower than 0.10 are displayed in boldface. The alternative hypothesis is too few runs.

Table B.12: Results of the Z_A test for equality of empirical distributions.

	Period				
Test	1994-1997	1997-2000	2000-2003	2003-2006	2006-2009
			Panel A: Unconditional Tests		
$\widehat{S}$	**0.01**	**0.00**	0.12	0.68	0.22
HS	0.10	**0.02**	0.67	**0.07**	**0.02**
HC	0.32	**0.00**	0.13	**0.00**	**0.05**
HP	0.16	**0.00**	0.64	**0.02**	0.54
KS	0.83	0.44	0.61	0.77	0.77
KC	1.00	**0.00**	0.83	0.22	0.47
KP	0.27	**0.00**	0.78	0.22	0.76
$\widehat{S}_\rho$	0.94	**0.00**	**0.01**	0.64	0.53
			Panel B: Conditional Tests		
$\widehat{\xi_i}$	**0.06**	**0.03**	**0.05**	0.31	0.10
CS	0.53	**0.00**	**0.05**	0.25	0.55
$\widehat{S}_\rho$	**0.05**	**0.00**	0.14	0.17	0.17

Note: Entries are the two-sided p-values of the Z_A test as defined in (2.25c) for the hypothesis (2.2), where p-values lower than 0.10 are displayed in boldface.

Table B.13: Results of the $\widehat{D}$ test for equality of density estimates

Test	Period 1994-1997	1997-2000	2000-2003	2003-2006	2006-2009
	Panel A: Unconditional Tests				
$\widehat{S}$	**0.01**	**0.01**	**0.06**	0.41	0.57
HS	**0.06**	**0.03**	0.43	**0.01**	0.19
HC	0.15	**0.04**	**0.06**	**0.01**	0.56
HP	0.22	**0.00**	**0.06**	**0.00**	**0.10**
KS	0.40	**0.05**	0.50	0.43	0.15
KC	0.68	**0.01**	0.24	0.21	**0.07**
KP	**0.04**	**0.00**	0.51	0.20	**0.07**
$\widehat{S}_\rho$	0.12	**0.02**	**0.03**	0.13	0.18
	Panel B: Conditional Tests				
$\widehat{\xi_i}$	**0.07**	**0.01**	**0.03**	0.16	**0.08**
CS	0.41	**0.00**	0.17	**0.09**	0.33
$\widehat{S}_\rho$	**0.01**	**0.00**	**0.02**	0.04	**0.01**

Note: Entries are the two-sided p-values of the $\widehat{D}$ test as defined in (2.26) for the hypothesis (2.2), where p-values lower than 0.10 are displayed in boldface.

Table B.14: Results of the Wilcoxon test for location shift

Test	Period 1994-1997	1997-2000	2000-2003	2003-2006	2006-2009
	Panel A: Unconditional Tests				
$\lvert\widehat{S}\rvert$	**0.03**	0.96	0.83	0.33	0.79
HS	0.91	0.97	0.90	0.65	0.91
HC	0.80	1.00	0.92	0.59	0.72
HP	0.92	1.00	0.90	0.69	0.82
KS	0.27	0.81	0.87	0.38	0.76
KC	0.50	1.00	0.85	0.29	**0.08**
KP	0.57	1.00	0.88	0.35	0.24
$\widehat{S}_\rho$	0.74	1.00	0.99	0.56	0.86
	Panel B: Conditional Tests				
$\lvert\widehat{\xi}_i\rvert$	**0.09**	0.92	0.77	0.06	0.95
CS	0.56	1.00	0.99	0.92	0.91
$\widehat{S}_\rho$	0.92	1.00	0.98	0.75	0.96

Note: Entries are one-sided p-values of the Wilcoxon rank-sum test for the hypothesis (2.4), where p-values lower than 0.10 are displayed in boldface.

Table B.15: Results of the tests for scale shift

Test	Period 1994-1997	1997-2000	2000-2003	2003-2006	2006-2009
	Panel A: Unconditional Symmetry				
Ansari-Bradley	**0.00**	**0.08**	0.16	0.77	0.79
Mood	**0.01**	0.11	**0.09**	0.76	0.91
	Panel B: Conditional Symmetry				
Ansari-Bradley	0.85	0.45	0.17	0.91	0.16
Mood	0.94	0.55	**0.06**	0.93	0.19

Note: Entries are the one-sided p-values of the Ansari-Bradley and Mood tests for the hypothesis (2.3). The unconditional asymmetry metric used is the skewness statistic $\widehat{S}(\eta_t)$ as proposed by Lisi, and conditional asymmetry is measured by the parameter $\widehat{\xi}_i$ of the GARCH innovation of returns. The results for the 1997-2000 period must be interpreted with caution, as $f_{S|R}(s|r=0)$ and $f_{S|R}(s|r=1)$ are likely to have a different location parameters during this period as suggested by the results in Table B.14. The distribution theory of the scale tests used is known only when the location of the two populations is the same or is known in advance and removed from the test samples.

Appendix C

C.1 Descriptive Statistics of Return Data

Table C.1: Measures of asymmetry in the returns of the market portfolio.

Period	Regression	$\hat{\xi}$
Panel A: Monthly Returns		
Full Period (2000 - 2009)	-3.16	0.65
	(0.0000)	(0.0001)
First Half (2000 - 2004)	-3.79	0.64
	(0.0005)	(0.0549)
Second Half (2005-2009)	-2.76	0.69
	(0.0064)	(0.0151)
Panel B: Weekly Returns		
Full Period (2000 - 2009)	-1.86	0.77
	(0.0906)	(0.0000)
First Half (2000 - 2004)	-4.17	0.80
	(0.0001)	(0.0063)
Second Half (2005-2009)	-1.24	0.75
	(0.2231)	(0.0007)

Notes: We used the centered series of excess returns of the Ibovespa from 2000:01 to 2009:12 in local currency as $\tilde{R}_m$ and the square of the same series as $\tilde{R}_m^2$. The column labeled "Regression" reports the slope of a regression of $\tilde{R}_m$ on $\tilde{R}_m^2$ and the associated p-value, which was obtained from Newey-West HAC standard errors. Under the assumption that returns follow a Skewed Student-t distribution à la Fernandez and Steel (1998), the second column reports maximum-likelihood estimates of the asymmetry parameter ξ and the associated two-sided p-value from the test $\xi = 1$. The distribution is symmetric when $\xi = 1$, and negatively (positively) skewed when ξ is lower (higher) than unity.

Table C.2: Summary statistics for size-sorted portfolios.

Portfolio	Mean	Std. Dev.	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\rho}_1 = \hat{\rho}_2 = 0$
Smallest	2.03	13.56	0.884***	-1.668*	-
2nd	-0.02	8.56	0.563***	-1.823*	-
3rd	1.47	9.22	0.659***	-2.672***	-
4th	0.10	8.81	0.718***	-2.417***	-
5th	-0.33	9.08	0.758***	-2.561***	-
6th	0.78	8.12	0.754***	-1.381*	0.97
7th	0.75	8.96	0.840***	-1.727***	-
8th	0.53	8.30	0.871***	-0.777*	-
9th	-0.22	7.60	0.811***	-0.705**	-
Largest	0.21	7.56	0.915***	0.365*	-

Notes: Data are monthly log returns of equally-weighted portfolios from 2000m01 to 2009m12 in excess of the risk-free rate (the CDI interbank rate). The mean and standard deviation of returns are expressed in percentages. The columns labeled "$\hat{\beta}_1$" and "$\hat{\beta}_2$" report the coefficients of the regression $r_{i,t} = \alpha_i + \beta_{1,i} r_{m,t} + \beta_{2,i} r_{m,t}^2 + \epsilon_{i,t}$ where $r_{i,t}$ and $r_{m,t}$ are the returns on portfolio i and on the broad market portfolio (Ibovespa) at time t. Significance levels are 10%(*), 5%(**) and 1%(***), based on Newey-West HAC standard errors with three lags. The last column concerns time-varying coskewness, and reports the overall significance (p-value) of a second-order autoregression for $u_{i,t} u_{m,t}^2$, where $u_{i,t}$ and $u_{m,t}^2$ are the residuals from regressing r_i and r_m^2 on r_m, respectively. $\hat{\rho}_j$ is the estimated coefficient of the j-th lagged term in the auto-regression, $u_{i,t-j} u_{m,t-j}^2$. A dash indicates a p-value of order 10^{-4} or less.

Table C.3: Summary statistics for portfolios sorted on price-to-book ratio.

Portfolio	Mean	Std. Dev.	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\rho}_1 = \hat{\rho}_2 = 0$
Lowest	1.14	12.50	0.845***	-1.303	-
2nd	0.45	9.64	0.852***	-1.515**	0.04
3rd	0.67	8.63	0.687***	-2.132***	-
4th	0.58	8.56	0.773***	-1.654***	-
5th	0.58	7.33	0.742***	-0.795*	0.75
6th	0.13	8.94	0.881***	-1.631***	-
7th	0.61	7.45	0.755***	-0.662*	0.94
8th	0.24	8.28	0.704***	-2.316***	-
9th	0.54	8.22	0.860***	-0.789**	-
Highest	0.42	8.64	0.705***	-2.307***	-

Notes: Data are monthly log returns of equally-weighted portfolios from 2000m01 to 2009m12 in excess of the risk-free rate (the CDI interbank rate). The mean and standard deviation of returns are expressed in percentages. The columns labeled "$\hat{\beta}_1$" and "$\hat{\beta}_2$" report the coefficients of the regression $r_{i,t} = \alpha_i + \beta_{1,i} r_{m,t} + \beta_{2,i} r_{m,t}^2 + \epsilon_{i,t}$ where $r_{i,t}$ and $r_{m,t}$ are the returns on portfolio i and on the broad market portfolio (Ibovespa) at time t. Significance levels are 10%(*), 5%(**) and 1%(***), based on Newey-West HAC standard errors with three lags. The last column concerns time-varying coskewness, and reports the overall significance (p-value) of a second-order autoregression for $u_{i,t} u_{m,t}^2$, where $u_{i,t}$ and $u_{m,t}^2$ are the residuals from regressing r_i and r_m^2 on r_m, respectively. $\hat{\rho}_j$ is the estimated coefficient of the j-th lagged term in the auto-regression, $u_{i,t-j} u_{m,t-j}^2$. A dash indicates a p-value of order 10^{-4} or less.

Table C.4: Summary statistics for industry portfolios

Portfolio	Mean	Std. Dev.	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\rho}_1 = \hat{\rho}_2 = 0$
Telecom	-0.73	8.63	0.894***	1.093**	-
Energy	0.37	8.06	0.790***	0.242	0.95
Food & Bev	0.35	10.16	0.773***	-1.470**	-
Retail	0.75	11.01	0.934***	-1.663***	0.01
Electronics	0.58	14.99	0.600***	-1.624	-
Banks	0.87	8.65	0.799***	-1.462***	-
Minerals	1.41	10.85	0.555***	-0.869	0.04
Mining	1.69	10.60	0.637***	-3.891***	-
Machinery	1.21	8.28	0.369***	-3.153***	-
Water	-0.29	10.09	0.802***	-0.986	0.03
Pulp & Paper	-0.14	9.64	0.612***	-2.821***	-
Oil & Gas	1.27	8.68	0.728***	-1.586*	0.01
Chemichals	0.67	8.73	0.644***	-1.768**	-
Steel	1.39	9.45	0.769***	-2.404***	-
Textiles	1.69	10.83	0.635***	-1.350*	-
Vehicles	1.79	9.44	0.559***	-2.826***	-
Transport	0.59	13.67	0.601**	-2.434***	-

See notes for Table C.2. Industry classifications as provided by the Economática database. The Agriculture, Construction and IT sectors did not have a meaningful number of companies with actively trading stocks during the whole 2000 - 2009 period, and were left out of the sample.

C.2 Estimation Results

Table C.5: Estimates of the KL/SW model with monthly returns

Estimate	Size Portfolios		PB Portfolios		Industry Portfolios	
	EW	VW	EW	VW	EW	VW
Mg. rate of subst.[a]	15.32	-5.76	14.12	-6.03	10.05	NA
	(0.000)	(0.997)	(0.000)	(0.734)	(0.011)	-
Market asymmetry[b]	1.82e-06	2.67e03	2.55e-06	-2.75e-06	3.16e-04	NA
	(0.003)	(0.997)	(0.000)	(0.584)	(0.000)	-
Coskewness loadings[c]	193.57	0.04	85.85	1.01	127.84	NA
	(0.000)	(0.997)	(0.000)	(0.999)	(0.000)	-
Hansen's J Test	9.62	7.98	10.53	5.66	6.27	NA
	(0.944)	(0.979)	(0.913)	(0.997)	(0.999)	-

Notes: The table reports two-step GMM estimates of selected parameters in the model (3.7), based on the moment conditions (3.8). In the second step, the weighting matrix is based on the Newey-West HAC estimator with three lags. The Hansen's J test reports the value of the statistic and the associated p-value. For size and price-to-book portfolios, the null distribution of J is $\chi^2(8)$, while for industry portfolios it is distributed as $\chi^2(15)$. Due to the high number of parameters ($3N + 4$, where $N = 10$ for the size and price-to-book portfolios and $N = 17$ for industry portfolios), only the full sample with $T = 120$ monthly returns from 2000m01 to 2009m12 was used. EW and VW stand for equally-weighted and value-weighted portfolios. The "NA" indicates failure of the GMM criterion function to converge.
[a] Estimate of ϕ and associated p-value of the test $\phi = 0$.
[b] Estimate of S_m and associated p-value of the test $S_m = 0$.
[c] Wald statistic of the joint test $\gamma_{2,i} = 0\,\forall i$ and associated p-value.

Table C.6: Estimates of the KL/SW model with weekly returns

Estimate	Full Sample EW	Full Sample VW	First Half EW	First Half VW	Second Half EW	Second Half VW
			Panel A: Size Portfolios			
Mg. rate of subst.[a]	4.93	9.74	NA	NA	-4.72	-0.92
	(0.682)	(0.419)	-	-	(0.979)	(0.173)
Market asymmetry[b]	6.36e-07	5.00e-07	NA	NA	1.97e-04	3.23e-06
	(0.686)	(0.405)	-	-	(0.987)	(0.438)
Coskewness loadings[c]	0.36	1.06	NA	NA	0.00	1.62
	(1.000)	(0.999)	-	-	(1.000)	(0.998)
Hansen's J Test	7.04	8.51	NA	NA	8.32	7.99
	(0.989)	(0.970)	-	-	(0.974)	(0.979)
			Panel B: Price-to-Book Portfolios			
Mg. rate of subst.[a]	5.68	0.45	NA	2.23	-1.22	NA
	(0.508)	(0.948)	-	(0.911)	(0.496)	-
Market asymmetry[b]	1.19e-06	5.17e-07	NA	2.34e-07	5.95e-06	NA
	(0.528)	(0.626)	-	(0.502)	(0.685)	-
Coskewness loadings[c]	0.55	0.30	NA	0.59	0.38	NA
	(1.000)	(1.000)	-	(1.000)	(1.000)	-
Hansen's J Test	6.39	3.41	NA	2.99	3.81	NA
	(0.994)	(0.999)	-	(1.000)	(0.999)	-
			Panel C: Industry Portfolios			
Mg. rate of subst.[a]	NA	NA	NA	NA	NA	NA
	-	-	-	-	-	-
Market asymmetry[b]	NA	NA	NA	NA	NA	NA
	-	-	-	-	-	-
Coskewness loadings[c]	NA	NA	NA	NA	NA	NA
	-	-	-	-	-	-
Hansen's J Test	NA	NA	NA	NA	NA	NA
	-	-	-	-	-	-

Notes: The table reports two-step GMM estimates of selected parameters in the model (3.7), based on the moment conditions (3.8). In the second step, the weighting matrix is based on the Newey-West HAC estimator with five lags. The Hansen's J test reports the value of the statistic and the associated p-value. For size and price-to-book portfolios, the null distribution of J is $\chi^2(8)$, while for industry portfolios it is distributed as $\chi^2(15)$. The full sample has $T = 520$ weekly returns from 2000w01 to 2009w12, while the first and second halves, from 2000w01 to 2004w12 and 2005w01 to 2009w12, have $T = 260$ each. EW and VW stand for equally-weighted and value-weighted portfolios. The "NA" indicates failure of the GMM criterion function to converge.

[a] Estimate of ϕ and associated p-value of the test $\phi = 0$.

[b] Estimate of S_m and associated p-value of the test $S_m = 0$.

[c] Wald statistic of the joint test $\gamma_{2,i} = 0\,\forall i$ and associated p-value.

Table C.7: Estimates of the arbitrage model with monthly returns

Estimate	Full Sample EW	Full Sample VW	First Half EW	First Half VW	Second Half EW	Second Half VW
			Panel A: Size Portfolios			
Market risk premium[a]	0.056	NA	0.135	NA	NA	0.515
	(0.229)	-	(0.705)	-	-	(0.893)
Coskewness premium[b]	0.018	NA	0.054	NA	NA	0.183
	(0.212)	-	(0.699)	-	-	(0.891)
Coskewness loadings[c]	76.36	NA	64.64	NA	NA	145.06
	(0.000)	-	(0.000)	-	-	(0.000)
Pricing errors[d]	50.28	NA	73.03	NA	NA	46.88
	(0.000)	-	(0.000)	-	-	(0.000)
Hansen's J Test	0.02	NA	2.56	NA	NA	0.00
	(1.000)	-	(0.959)	-	-	(0.000)
			Panel B: Price-to-Book Portfolios			
Market risk premium[a]	0.074	0.028	0.027	NA	NA	NA
	(0.799)	(0.141)	(0.276)	-	-	-
Coskewness premium[b]	0.039	0.010	0.007	NA	NA	NA
	(0.801)	(0.235)	(0.445)	-	-	-
Coskewness loadings[c]	117.74	389.97	13.02	NA	NA	NA
	(0.000)	(0.000)	(0.223)	-	-	-
Pricing errors[d]	39.57	27.19	40.09	NA	NA	NA
	(0.000)	(0.002)	(0.000)	-	-	-
Hansen's J Test	0.00	1.74	0.00	NA	NA	NA
	(0.000)	(0.988)	(0.000)	-	-	-
			Panel C: Industry Portfolios			
Market risk premium[a]	NA	NA	NA	NA	0.173	NA
	-	-	-	-	(0.843)	-
Coskewness premium[b]	NA	NA	NA	NA	0.061	NA
	-	-	-	-	(0.853)	-
Coskewness loadings[c]	NA	NA	NA	NA	80.47	NA
	-	-	-	-	(0.000)	-
Pricing errors[d]	NA	NA	NA	NA	34.29	NA
	-	-	-	-	(0.000)	-
Hansen's J Test	NA	NA	NA	NA	0.00	NA
	-	-	-	-	(1.000)	-

Notes: The table reports two-step GMM estimates of selected parameters in the model (3.12)-(3.13), based on the moment conditions (3.14). In the second step, the weighting matrix is based on the Newey-West HAC estimator with three lags. The Hansen's J test reports the value of the statistic and the associated p-value. For size and price-to-book portfolios, the null distribution of J is $\chi^2(8)$, while for industry portfolios it is distributed as $\chi^2(15)$. The full sample has $T = 120$ monthly returns from 2000m01 to 2009m12, while the first and second halves, from 2000m01 to 2004m12 and 2005m01 to 2009m12, have $T = 60$ each. EW and VW stand for equally-weighted and value-weighted portfolios. The "NA" indicates failure of the GMM criterion function to converge.

[a] Estimate of λ_1 and associated p-value of the test $\lambda_1 = 0$.

[b] Estimate of λ_2 and associated p-value of the test $\lambda_2 = 0$.

[c] Wald statistic for the joint test of nullity of the coskewness coefficients, $\beta_{2,i} = 0\,\forall i$.

[d] Wald statistic for the joint test of nullity of the pricing errors, $\beta_{0,i} = 0\,\forall i$.

Table C.8: Estimates of the arbitrage model with weekly returns

Estimate	Full Sample EW	Full Sample VW	First Half EW	First Half VW	Second Half EW	Second Half VW
			Panel A: Size Portfolios			
Market risk premium[a]	0.010	NA	0.009	0.010	NA	NA
	(0.081)	-	(0.031)	(0.536)	-	-
Coskewness premium[b]	0.004	NA	0.006	-0.045	NA	NA
	(0.048)	-	(0.069)	(0.560)	-	-
Coskewness loadings[c]	72.94	NA	95.50	145.02	NA	NA
	(0.000)	-	(0.000)	(0.000)	-	-
Pricing errors[d]	81.95	NA	90.64	38.37	NA	NA
	(0.000)	-	(0.000)	(0.000)	-	-
Hansen's J Test	3.28	NA	1.33	0.00	NA	NA
	(0.916)	-	(0.995)	(1.000)	-	-
			Panel B: Price-to-Book Portfolios			
Market risk premium[a]	NA	NA	NA	NA	0.076	0.008
	-	-	-	-	(0.846)	(0.183)
Coskewness premium[b]	NA	NA	NA	NA	0.035	0.002
	-	-	-	-	(0.854)	(0.226)
Coskewness loadings[c]	NA	NA	NA	NA	101.72	52.37
	-	-	-	-	(0.000)	(0.000)
Pricing errors[d]	NA	NA	NA	NA	28.38	19.57
	-	-	-	-	(0.002)	(0.034)
Hansen's J Test	NA	NA	NA	NA	0.00	0.00
	-	-	-	-	(1.000)	(1.000)
			Panel C: Industry Portfolios			
Market risk premium[a]	NA	NA	0.134	NA	NA	NA
	-	-	(0.584)	-	-	-
Coskewness premium[b]	NA	NA	0.047	NA	NA	NA
	-	-	(0.604)	-	-	-
Coskewness loadings[c]	NA	NA	269.02	NA	NA	NA
	-	-	(0.000)	-	-	-
Pricing errors[d]	NA	NA	34.98	NA	NA	NA
	-	-	(0.000)	-	-	-
Hansen's J Test	NA	NA	0.00	NA	NA	NA
	-	-	(1.000)	-	-	-

Notes: The table reports two-step GMM estimates of selected parameters in the model (3.12)-(3.13), based on the moment conditions (3.14). In the second step, the weighting matrix is based on the Newey-West HAC estimator with five lags. The Hansen's J test reports the value of the statistic and the associated p-value. For size and price-to-book portfolios, the null distribution of J is $\chi^2(8)$, while for industry portfolios it is distributed as $\chi^2(15)$. The full sample has $T = 520$ weekly returns from 2000w01 to 2009w12, while the first and second halves, from 2000w01 to 2004w12 and 2005w01 to 2009w12, have $T = 260$ each. EW and VW stand for equally-weighted and value-weighted portfolios. The "NA" indicates failure of the GMM criterion function to converge.

[a] Estimate of λ_1 and associated p-value of the test $\lambda_1 = 0$.

[b] Estimate of λ_2 and associated p-value of the test $\lambda_2 = 0$.

[c] Wald statistic for the joint test of nullity of the coskewness coefficients, $\beta_{2,i} = 0\,\forall i$.

[d] Wald statistic for the joint test of nullity of the pricing errors, $\beta_{0,i} = 0\,\forall i$.

Table C.9: Estimates of the quadratic SDF with monthly returns

Estimate	Full Sample EW	Full Sample VW	First Half EW	First Half VW	Second Half EW	Second Half VW
			Panel A: Size Portfolios			
Covariance [a]	-2.39	-2.25	-2.09	-0.74	-6.72	-6.48
	(0.027)	(0.023)	(0.027)	(0.473)	(0.000)	(0.000)
Coskewness [b]	-63.44	-41.59	-114.24	-28.09	-49.38	-53.85
	(0.000)	(0.005)	(0.000)	(0.066)	(0.000)	(0.000)
Hansen's J Test	13.96	15.58	8.87	9.17	8.20	10.06
	(0.083)	(0.049)	(0.353)	(0.329)	(0.414)	(0.261)
			Panel B: Price-to-Book Portfolios			
Covariance [a]	-1.04	-1.37	-0.91	-0.29	-4.61	-5.29
	(0.390)	(0.158)	(0.293)	(0.787)	(0.001)	(0.000)
Coskewness [b]	-15.03	-14.61	-54.11	-20.99	-29.61	-32.38
	(0.224)	(0.084)	(0.002)	(0.064)	(0.000)	(0.000)
Hansen's J Test	4.24	8.86	5.75	7.89	4.34	6.76
	(0.835)	(0.355)	(0.675)	(0.444)	(0.825)	(0.563)
			Panel C: Industry Portfolios			
Covariance [a]	-2.13	-0.72	-1.34	0.27	-6.19	-7.03
	(0.101)	(0.443)	(0.020)	(0.725)	(0.000)	(0.000)
Coskewness [b]	-18.61	-5.96	-102.63	-33.57	-38.28	-62.03
	(0.134)	(0.329)	(0.000)	(0.035)	(0.000)	(0.000)
Hansen's J Test	18.54	16.38	10.81	11.39	9.48	10.23
	(0.236)	(0.357)	(0.767)	(0.724)	(0.851)	(0.805)

Notes: The table reports two-step GMM estimates of θ_1 and θ_2 as specified in (3.15), with θ_0 normalized to 1 and based on the moment conditions (3.16). In the second step, the weighting matrix is based on the Newey-West HAC estimator with three lags. The Hansen's J test reports the value of the statistic and the associated p-value. For size and price-to-book portfolios, the null distribution of J is $\chi^2(8)$, while for industry portfolios it is distributed as $\chi^2(15)$. The full sample has $T = 120$ monthly returns from 2000m01 to 2009m12, while the first and second halves, from 2000m01 to 2004m12 and 2005m01 to 2009m12, have $T = 60$ each. EW and VW stand for equally-weighted and value-weighted portfolios.

[a] Estimate of θ_1 and p-value of the test $\theta_1 = 0$.

[b] Estimate of θ_2 and p-value of the test $\theta_2 = 0$.

Table C.10: Estimates of the quadratic SDF with weekly returns

Estimate	Full Sample		First Half		Second Half	
	EW	VW	EW	VW	EW	VW
			Panel A: Size Portfolios			
Covariance [a]	-0.26	-0.65	-1.47	-1.30	-1.24	-1.62
	(0.763)	(0.490)	(0.371)	(0.424)	(0.197)	(0.126)
Coskewness [b]	-8.30	1.35	-27.70	28.26	-89.04	-106.27
	(0.727)	(0.967)	(0.452)	(0.507)	(0.011)	(0.014)
Hansen's J Test	21.44	15.89	12.28	8.50	13.08	16.18
	(0.006)	(0.044)	(0.139)	(0.386)	(0.109)	(0.040)
			Panel B: Price-to-Book Portfolios			
Covariance [a]	-1.11	-1.35	0.06	-1.07	-2.53	-2.64
	(0.319)	(0.233)	(0.965)	(0.507)	(0.085)	(0.068)
Coskewness [b]	1.98	12.56	-46.00	19.64	-73.66	-75.69
	(0.977)	(0.795)	(0.382)	(0.654)	(0.013)	(0.006)
Hansen's J Test	7.25	8.32	10.60	10.96	4.38	4.15
	(0.510)	(0.403)	(0.226)	(0.204)	(0.822)	(0.843)
			Panel C: Industry Portfolios			
Covariance [a]	-0.17	-0.06	-0.12	1.25	-0.81	-1.62
	(0.830)	(0.929)	(0.928)	(0.145)	(0.511)	(0.108)
Coskewness [b]	-36.48	-16.11	-111.62	-51.22	-44.52	-112.94
	(0.168)	(0.384)	(0.042)	(0.071)	(0.317)	(0.010)
Hansen's J Test	24.40	15.54	21.72	22.09	13.79	13.73
	(0.059)	(0.413)	(0.115)	(0.105)	(0.541)	(0.546)

Notes: The table reports two-step GMM estimates of θ_1 and θ_2 as specified in (3.15), with θ_0 normalized to 1 and based on the moment conditions (3.16). In the second step, the weighting matrix is based on the Newey-West HAC estimator with five lags. The Hansen's J test reports the value of the statistic and the associated p-value. For size and price-to-book portfolios, the null distribution of J is $\chi^2(8)$, while for industry portfolios it is distributed as $\chi^2(15)$. The full sample has $T = 520$ weekly returns from 2000w01 to 2009w52, while the first and second halves, from 2000w01 to 2004w52 and 2005w01 to 2009w52, have $T = 260$ each. EW and VW stand for equally-weighted and value-weighted portfolios.

[a] Estimate of θ_1 and p-value of the test $\theta_1 = 0$.

[b] Estimate of θ_2 and p-value of the test $\theta_2 = 0$.

Bibliography

Bibliography

[1] ALEXANDER, I., ESTACHE, A., AND OLIVERI, A. A few things transport regulators should know about risk and the cost of capital. *Utilities Policy 9*, 1 (2000), 1 – 13.

[2] ALEXANDER, I., MAYER, C., AND WEEDS, H. Regulatory structure and risk and infrastructure firms: an international comparison. Policy Research Working Paper 1698, The World Bank, 1996.

[3] ANDERSON, G. The power and size of nonparametric tests for common distributional characteristics. *Econometric reviews 20*, 1 (2001), 1–30.

[4] ARMSTRONG, M., AND SAPPINGTON, D. E. Recent developments in the theory of regulation. In *Handbook of Industrial Organization, Vol. 3*, M. Armstrong and R. Porter, Eds. Elsevier, 2007, ch. 27, pp. 1557 – 1700.

[5] ARROW, K. *Essays in the Theory of Risk-Bearing.* Chicago: Markham Publishing Co., 1971.

[6] ATHAYDE, G. M., AND FLÔRES, R. G. Introducing higher moments in the CAPM: Some basic ideas. In *Advances in Quantitative Asset Management*, C. L. Dunis, Ed. Kluwer Academic Publishers, 2000, ch. 1, pp. 3–15.

[7] ATHAYDE, G. M., AND FLÔRES, R. G. Finding a maximum skewness portfolio–a general solution to three-moments portfolio choice. *Journal of Economic Dynamics and Control 28*, 7 (2004), 1335 – 1352.

[8] BAI, J., AND NG, S. A consistent test for conditional symmetry in time series models. *Journal of Econometrics 103*, 1-2 (2001), 225–258.

[9] BAI, J., AND NG, S. Tests for skewness, kurtosis and normality for time series data. *Journal of Business and Economic Statistics 23*, 1 (2005), 49–60.

[10] BALLARD, C. L., SHOVEN, J. B., AND WHALLEY, J. General equilibrium computations of the marginal welfare costs of taxes in the United States. *The American Economic Review 75*, 1 (1985), 128–138.

[11] BARON, D. P., AND MYERSON, R. B. Regulating a monopolist with unknown costs. *Econometrica 50*, 4 (1982), 911–930.

[12] BARONE-ADESI, G. Arbitrage equilibrium with skewed asset returns. *The Journal of Financial and Quantitative Analysis 20*, 3 (1985), 299–313.

[13] BARONE-ADESI, G., GAGLIARDINI, P., AND URGA, G. Testing asset pricing models with coskewness. *Journal of Business and Economic Statistics 22*, 4 (2004), 474–485.

[14] Binder, J. J. Measuring the effects of regulation with stock price data. *The RAND Journal of Economics 16*, 2 (1985), 167–183.

[15] Black, F., Jensen, M. C., and Scholes, M. S. The Capital Asset Pricing Model: Some empirical tests. In *Studies in the Theory of Capital Markets*, M. Jensen, Ed. Praeger Publishers Inc., NY, 1972.

[16] Bohi, D. R., and Zimmerman, M. B. An update on econometric studies of energy demand behavior. *Annual Review of Energy 9*, 1 (1984), 105–154.

[17] Bowley, A. L. *Elements of Statistics*, 6th ed. New York: Scribner, 1937.

[18] Brennan, M. J., and Schwartz, E. S. Consistent regulatory policy under uncertainty. *The Bell Journal of Economics 13*, 2 (1982), 506–521.

[19] Castro, Jr., F. H. *Apreçamento de Ativos com Assimetria e Curtose: Um Teste de Comomentos com Dados em Painel.* PhD thesis, Universidade de São Paulo, 2008.

[20] Chen, Y. T., and Lin, C. C. On the robustness of symmetry tests for stock returns. *Studies in Nonlinear Dynamics & Econometrics 12*, 2 (2008), 1591–1631.

[21] Cochrane, J. H. *Asset Pricing.* Princeton University Press, 2005.

[22] Cramér, H. *Mathematical Methods of Statistics.* Princeton University Press, Princeton, 1946.

[23] Crémer, J., and McLean, R. P. Full extraction of the surplus in bayesian and dominant strategy auctions. *Econometrica 56*, 6 (1988), 1247–57.

[24] da Silva, A. C. Modeling and estimating a higher systematic co-moment asset pricing model in the brazilian stock market. *Latin American Business Review 6*, 4 (2004), 85–101.

[25] Davidson, W. N., Rangan, N., and Rosenstein, S. Regulation and systematic risk in the electric utility industry: A test of the buffering hypothesis. *The Financial Review 32*, 1 (February 1997), 163–84.

[26] de Lima, P. J. F. On the robustness of nonlinearity tests to moment condition failure. *Journal of Econometrics 76*, 1-2 (1997), 251–280.

[27] Demougin, D. M., and Garvie, D. A. Contractual design with correlated information under limited liability. *RAND Journal of Economics 22*, 4 (1991), 477 – 489.

[28] Ding, Z., Granger, C. W. J., and Engle, R. F. A long memory property of stock market returns and a new model. *Journal of Empirical Finance 1*, 1 (1993), 83–106.

[29] Dittmar, R. F. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *The Journal of Finance 57*, 1 (2002), 369–403.

[30] Fama, E. F., and MacBeth, J. D. Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy 81*, 3 (1973), 607–636.

[31] Fernandez, C., and Steel, M. F. J. On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association 93*, 441 (1998).

[32] Fisher, R. A. On the interpretation of χ^2 from contingency tables, and the calculation of P. *Journal of the Royal Statistical Society 85*, 1 (1922), 87–94.

[33] Friend, I., and Westerfield, R. Co-skewness and capital asset pricing. *The Journal of Finance 35*, 4 (1980), 897–913.

[34] Fuller, W. A. *Introduction to Statistical Time Series.* John Wiley & Sons, Inc., 1976.

[35] Gallant, A. R. Seemingly unrelated nonlinear regressions. *Journal of Econometrics 3*, 1 (1975), 35 – 50.

[36] Gasmi, F., Ivaldi, M., and Laffont, J.-J. Rent extraction and incentives for efficiency in recent regulatory proposals. *Journal of Regulatory Economics 6*, 2 (1994), 151–176.

[37] Gasmi, F., Kennet, D. M., Laffont, J.-J., and Sharkey, W. W. *Cost Proxy Models And Telecommunications Policy.* MIT Press, 2002.

[38] Gibbons, J. D., and Chakraborti, S. *Nonparametric Statistical Inference.* Marcel Dekker, 2003.

[39] Gibbons, M. R. *Econometric Methods for Testing a Class of Financial Models.* PhD thesis, University of Chicago, 1980.

[40] Gibbons, M. R. Multivariate tests of financial models: A new approach. *Journal of Financial Economics 10*, 1 (1982), 3 – 27.

[41] Glosten, L. R., Jagannathan, R., and Runkle, D. E. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance 48*, 5 (1993), 1779–1801.

[42] Granger, C. W., Maasoumi, E., and Racine, J. A dependence metric for possibly nonlinear processes. *Journal of Time Series Analysis 25*, 5 (2004), 649–669.

[43] Grout, P. A., and Zalewska, A. The impact of regulation on market risk. *Journal of Financial Economics 80*, 1 (2006), 149 – 184.

[44] Guthrie, G. Regulating infrastructure: The impact on risk and investment. *Journal of Economic Literature 44*, 4 (2006), 925–972.

[45] Harvey, C. R., and Siddique, A. Conditional skewness in asset pricing tests. *The Journal of Finance 55*, 3 (2000), 1263–1295.

[46] Hausman, J. A., and Poterba, J. M. Household behavior and the tax reform act of 1986. *The Journal of Economic Perspectives 1*, 1 (1987), 101–119.

[47] Hinkley, D. V. On power transformations to symmetry. *Biometrika 62*, 1 (1975), 101–111.

[48] Ibbotson. *Stocks, Bonds, Bills and Inflation, 2007 Yearbook.* Morningstar, Inc., Chicago, IL., 2007.

[49] Jansen, D. W., and De Vries, C. G. On the frequency of large stock returns: putting booms and busts into perspective. *The Review of Economics and Statistics 73*, 1 (1991), 18–24.

[50] Jean, W. H. The extension of portfolio analysis to three or more parameters. *The Journal of Financial and Quantitative Analysis 6*, 1 (1971), 505–515.

[51] Jean, W. H. More on multidimensional portfolio analysis. *The Journal of Financial and Quantitative Analysis 8*, 3 (1973), 475–490.

[52] Jenkinson, T. Regulation and the Cost of Capital. In *International Handbook on Economic Regulation*, M. Krew and D. Parker, Eds. Edward Elgar Publishing Limited, 2006.

[53] Joanes, D. N., and Gill, C. A. Comparing measures of sample skewness and kurtosis. *The Statistician 47*, 1 (1998), 183–189.

[54] Kolbe, A. L., Tye, W. B., and Myers, S. *Regulatory Risk: Economic Principles and Applications to Natural Gas Pipelines and Other Industries.* Springer, 1993.

[55] Kraus, A., and Litzenberger, R. H. Skewness preference and the valuation of risk assets. *The Journal of Finance 31*, 4 (1976), 1085–1100.

[56] Laffont, J.-J., and Tirole, J. Using cost observation to regulate firms. *The Journal of Political Economy 94*, 3 (1986), 614–641.

[57] Laffont, J.-J., and Tirole, J. *A Theory of Incentives in Procurement and Regulation.* MIT Press, 1993.

[58] Lee, A., Moy, R. L., and Lee, C. F. A multivariate test of the covariance-co-skewness restriction for the three moment CAPM. *Journal of Economics and Business 48*, 5 (1996), 515 – 523.

[59] Li, Q., and Racine, J. S. *Nonparametric Econometrics: Theory and Practice.* Princeton University Press, Princeton, NJ, 2007.

[60] Lim, K.-G. A new test of the three-moment capital asset pricing model. *The Journal of Financial and Quantitative Analysis 24*, 2 (1989), 205–216.

[61] Lisi, F. Testing asymmetry in financial time series. *Quantitative Finance 7*, 6 (2007), 687–696.

[62] Lomnicki, Z. A. Test for departure from normality in the case of linear stochastic processes. *Metrika 4* (1961), 37–62.

[63] Loretan, M., and Phillips, P. C. B. Testing the covariance stationarity of heavy-tailed time series: An overview of the theory with applications to several financial datasets. *Journal of Empirical Finance 1*, 2 (1994), 211 – 248.

[64] Maasoumi, E., and Racine, J. Entropy andpred ictability of stock market entropy andpred ictability of stock market returns. *Journal of Econometrics 107*, 2 (2002), 291–312.

[65] Maasoumi, E., and Racine, J. A robust entropy-based test of asymmetry for discrete and continuous processes. *Econometric reviews 28*, 1-3 (2009), 246–261.

[66] Marshall, W. J., Yawitz, J. B., and Greenberg, E. Optimal regulation under uncertainty. *The Journal of Finance 36*, 4 (1981), 909–921.

[67] Moors, J., Wagemakers, R., Coenen, V., Heuts, R., and Janssens, M. Characterizing systems of distributions by quantile measures. *Statistica Neerlandica 50*, 3 (1996), 417 – 430.

[68] Myers, S. C. The application of finance theory to public utility rate cases. *The Bell Journal of Economics and Management Science 3*, 1 (1972), 58–97.

[69] Ngatchou-Wandji, J. On testing for the nullity of some skewness coefficients. *International Statistical Review 74*, 1 (2006), 47–65.

[70] NORTON, S. W. Regulation and systematic risk: The case of electric utilities. *Journal of Law and Economics 28*, 3 (1985), 671–686.

[71] PELTZMAN, S. Toward a more general theory of regulation. *Journal of Law and Economics 19*, 2 (1976), 211–40.

[72] POLITIS, D. N., AND ROMANO, J. P. The stationary bootstrap. *Journal of the American Statistical Association 89*, 428 (1994), 1303–1313.

[73] RIORDAN, M. H., AND SAPPINGTON, D. E. M. Optimal contracts with public ex-post information. *Journal of Economic Theory 45*, 1 (1988), 189 – 199.

[74] RUBINSTEIN, M. E. The fundamental theorem of parameter-preference security valuation. *The Journal of Financial and Quantitative Analysis 8*, 1 (1973), 61–69.

[75] SAMUELSON, P. A. The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. *The Review of Economic Studies 37*, 4 (1970), 537–542.

[76] SCHMALENSEE, R. Good regulatory regimes. *The RAND Journal of Economics 20*, 3 (1989), 417–436.

[77] SCHWERT, G. W. Using financial data to measure effects of regulation. *Journal of Law and Economics 24*, 1 (1981), 121–158.

[78] SEARS, R. S., AND WEI, K. C. Asset pricing, higher moments, and the market risk premium: A note. *The Journal of Finance 40*, 4 (1985), 1251–1253.

[79] SEARS, R. S., AND WEI, K. C. The structure of skewness preferences in asset pricing models with higher moments: an empirical test. *The Financial Review 23*, 1 (1988), 25–38.

[80] SHANKEN, J. On the estimation of beta pricing models. *Review of Financial Studies 5*, 1-34 (1992).

[81] SINGLETON, J. C., AND WINGENDER, J. Skewness persistence in common stock returns. *Journal of Financial and Quantitative Analisys 21*, 3 (1986), 335–341.

[82] SMITH, D. R. Conditional coskewness and asset pricing. *Journal of Empirical Finance 14*, 1 (2007), 91 – 119.

[83] STIGLER, G. J. The theory of economic regulation. *Bell Journal of Economics 2*, 1 (1971), 3–21.

[84] STUART, A., AND ORD, J. *Kendall's Advanced Theory of Statistics.* Oxford University Press, 1987.

[85] TAYLOR, L. D. The demand for electricity: A survey. *The Bell Journal of Economics 6*, 1 (1975), 74–110.

[86] TAYLOR, L. D. *Telecommunications Demand in Theory and Practice.* Kluwer Academic Publishers, 1994.

[87] TERÄSVIRTA, T. An Introduction to Univariate GARCH Models. In *Handbook of Financial Time Series*, T. G. Andersen, R. A. Davis, J.-P. Kreiss, and T. Mikosch, Eds. Springer Berlin Heidelberg, 2009.

[88] TSAY, R. S. *Analysis of Financial Time Series*, 2nd ed. John Wiley and Sons, 2005.

[89] van Zwet, W. R. *Convex Transformations of Random Variables.* Mathematisch Centrum, Amsterdam, 1964.

[90] Viscusi, W. K., Vernon, J., and Harrington, J. E. *Economics of Regulation and Antitrust*, 4th ed. MIT Press, 2005.

[91] Wright, S., Mason, R., and Miles, D. A study into certain aspects of the cost of capital for regulated utilities in the U.K. Tech. rep., Smithers & Co Ltd, London, 2003.

[92] Zhang, J. Powerful two-sample tests based on the likelihood ratio. *Technometrics 48*, 1 (2006), 95–103.