UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Systemic Risk and Returns

Permalink
https://escholarship.org/uc/item/4m10k9cj

Author
Boman, Arthur Lee

Publication Date
2013

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Systemic Risk and Returns

by

Arthur Lee Boman

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Peter Berck, Chair
Professor Shachar Kariv
Professor Jean Helwege
Professor Larry Karp

Fall 2013

Abstract

Systemic Risk and Returns

by

Arthur Lee Boman

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Peter Berck, Chair

I solve a consumption based model, with interfirm systemic risk, for a portfolio optimization with arbitrary return distributions and endogenous stochastic discount factor (sdf). The model highlights a new systemic risk: systemic allocation risk. In contrast to the case without systemic risk, the market and planner allocate capital differently. The externality causes the planner to reduce investment in the risky firm. The market, modeled as a representative agent, does not just ignore the externality and invest as if there were none. Instead, systemic risk increases the representative agent's investment in the systemically risky institution or industry, further increasing systemic risk. I introduce bailout of the financial industry and find it has a beneficial direct effect and a distortion effect. In some cases, *investor* moral hazard can make *ex post* optimal bailouts reduce *ex ante* utility - even when bailout does not benefit the financial industry's investors. I show that systemic risk, as opposed to systematic risk, can be characterized as a situation where the fundamental theorems of asset pricing do not apply.

Next I put the the model into a factor model, using the arbitrage pricing theory for market pricing of the firms. I use the model to distinguish between systematic and systemic risks. By directly including systemic risk, the potential of an interfirm or inter-industry externality, the model shows that including terms with fat tails in specifications for returns does not make them systemic risk if they still meet the definition of systematic risk (Systematic risk is risk within a firm's returns that is both non-causal and correlated with the stochastic discount factor - and therefore undiversifiable. In a factor model, systematic risk in the financial industry is the overall magnitude of firms loading onto systematic factors. The systematic factors do not need to be Gaussian.). The model shows why systematic risk is so often mistaken as systemic risk, why systematic risk in the financial industry is important, and why it should be considered along with systemic risk in regulatory efforts. The model is then used to delineate and outline the various

types of risk. This vocabulary can facilitate communication and research in systemic risk. Finally, I derive a popular systemic risk measure directly in terms of the parameters of a pricing model. I test the one that attempts to include causality in its measure, $\Delta CoVaR$.

$\Delta CoVaR$ seeks to use joint return data to measure a firm's contribution to systemic risk. To learn what comprehensive regulatory changes can do to systemic risk in general, and $\Delta CoVaR$ in particular, Part 4 estimates the impact of the extensive and coincident U.S. regulatory changes of 1993 (including Prompt Corrective Action law and Basel I) on the systemic risk level of commercial banks, as measured by $\Delta CoVaR$. Investment banks not subject to the law are used as controls. In a difference-in-difference framework, the law is used as a treatment shock. Use of a novel $\Delta CoVaR$ measure (unconditional rolling $\Delta CoVaR$) allows econometric assessment of exogenous changes and estimation of $\Delta CoVaR$ standard errors. With high power, no effect is found. This eliminates from possibility one of two formerly widely held beliefs that are each the basis of a literature: 1. That PCA and concurrent regulation lowered systemic risk, or 2. That $\Delta CoVaR$ measures systemic risk. The unique circumstances used for this test could also be exploited to assess other systemic risk metrics or inform other risk/regulation questions.

This PhD is dedicated to my parents, Cecille, and my whole family.

This dissertation is dedicated to the learning of finance wherever it may occur.

Contents

1 Systemic Risk with an Endogenous Stochastic Discount Factor **1**
- 1.1 Introduction . . . 1
- 1.2 Model and Pricing . . . 5
  - 1.2.1 Notation, Assumptions, Discussion of Assumptions . . . 5
  - 1.2.2 The Planner's Optimization . . . 7
  - 1.2.3 The Representative Agent's Optimization . . . 9
- 1.3 Mitigating Systemic Risk with Bailouts . . . 11
  - 1.3.1 The General Case . . . 11
  - 1.3.2 Two-industry Model . . . 13
- 1.4 Part 1 Conclusion . . . 15
- 1.5 Part 1 Proofs . . . 16

2 Equilibrium **20**
- 2.1 Epistemology . . . 20
- 2.2 Existence and Uniqueness . . . 22
  - 2.2.1 Set-up . . . 22
  - 2.2.2 Social Planner . . . 24
  - 2.2.3 Aggregate Agent . . . 25
- 2.3 Part 2 Conclusion . . . 26
- 2.4 Part 2 Proofs . . . 27

3 Arbitrage Pricing Theory and Systemic Risk **30**
- 3.1 Introduction . . . 30
- 3.2 Assumptions and Model . . . 31
  - 3.2.1 Basic Pricing of the Systemically Impacted Firms . . . 32
  - 3.2.2 Arbitrage Pricing of the Systemically Impacted Firms . . . 32
- 3.3 Application: *The Various Types of Systemic Risk* . . . 34
- 3.4 Potential Applications for Returns . . . 39
- 3.5 Derivation of Systemic Risk Measure from an Asset Pricing Model . . . 40
- 3.6 Part 3 Conclusion . . . 43

4 Measuring Systemic Risk **45**
- 4.1 Introduction . . . 45
- 4.2 The Two Strands of Literature . . . 47
  - 4.2.1 PCA and Concurrent Changes . . . 47
  - 4.2.2 **ΔCoVaR**, a Measure of a Firm's Contribution to Systemic Risk . . . 48
- 4.3 Data, the Two Groups, and the Treatment . . . 49
- 4.4 Estimating the Dependent Variable . . . 52
  - 4.4.1 Which **ΔCoVaR**? . . . 52
  - 4.4.2 Estimating Unconditional Rolling **ΔCoVaR**'s . . . 54
- 4.5 Treatment Effect Estimate . . . 56

4.6 Results 57
4.7 Results Discussion 59
4.8 **ΔCoVaR** Decomposition 60
4.9 Part 4 Conclusion 61

I would like to thank Peter Berck, Shachar Kariv, Leo Simon, Jean Helwege, Allen Berger, Larry Karp, Bob Anderson, an anonymous Financial Management Association (FMA) Conference reviewer, and participants at the 2012 FMA Conference for invaluable guidance and comments. I would like to thank Peter Berck for helping me to grow personally. I would like to thank the Agricultural and Resource Economics for research assistance funding and other funding.

1 Systemic Risk with an Endogenous Stochastic Discount Factor

1.1 Introduction

When successfully distinguished from systematic risk, systemic risk is about *causality.* Firms in the financial industry have linkages and connections which can become pathways to spread poor performance from one firm to the others - I define this phenomenon as *interfirm systemic risk* (ISR). When analyzing systemic risk, a vast literature focuses on the spreading of risk or distress in a causal way between firms within the financial industry. For well over two decades, papers have been published which focus on various mechanisms for the externality, such as counterparty risk and margin-spiral risk Humphrey (1986); Rochet and Tirole (1996); Cohen and Roberds (1993); Angelini et al. (1996). The risk of some type of causal externality (interfirm systemic risk) is an indispensable part of the story of systemic risk: Rochet and Tirole (1996); Gray and Jobst (2010); Angelini et al. (1996); Longstaff (2010); Bai et al. (2012); Adrian and Brunnermeier (2011). In contrast to systematic risk, these impacts are viewed as giving rise to more than just correlation. But how do we distinguish between correlation and causality, in their general forms, in an asset pricing model? How can we include ISR in the canonical consumption-based asset pricing model, to see the fully endogenous response of the representative agent?

One of the most common views of systemic risk is poor performance in one firm or group of firms negatively affecting other firms as spillovers or contagion, etc. Regardless of the specific mechanism, this can often be modeled as low returns in one firm having a negative impact on the returns of another firm. With the reasonable assumption that the impact on the other firm is in percentage terms, the impact becomes a direct reduction of the returns of the firm being affected. This allows everything to be written in returns, which in turn allows asset pricing models and methods.

There is another externality often called systemic risk. During a financial crisis (often just modeled as a highly negative return realization of the industry as a whole), the resulting impairment of the intermediation capacity of the financial industry is a negative externality affecting the real economy, such as in Acharya et al. (2010); Ibragimov et al. (2011); Huang et al. (2009) and Allen et al. (2012)). This externality is also a systemic risk, that I will call inter-industry systemic risk. What implication does this externality have on the size of industries or the returns demanded from certain industries?

Many papers have modeled particular mechanisms for interfirm systemic risk (ISR), and in such cases the model is specific to the particular mechanism and has not always fit well[1] into our pricing models, and the ones that have fit are (by design) specific. Many prior models of particular mechanisms of ISR can be reduced to a functional form of the returns of the firm or firms causing the externality - counterparty risk would be a step function for example. Of course, other researchers have added general externalities, or solved for the type of externality that would lead to equilibrium, rather than model a

[1] See Bai et al. (2012), or Gray et al. (2007) for exceptions.

mechanism causing the externality.

The key difference between my work and prior work is that I remain agnostic to the functional form of ISR, and that I model ISR in returns space without introducing other primitives. (It should also be clear that the externality is not the revelation or diffusion of information as in Andersen and Bollerslev (1997), information cascades as in Welch (1992); Grenadier (1996), or related to hidden information, for example the frailty of Duffie et al. (2009). I am not modeling networks as in Allen and Gale (2000) and Gale and Kariv (2007), or trade.)

Since at least Diamond (1967), it has been known that the market, modeled as a representative agent, will achieve constrained Pareto optimal capital allocation in the canonical consumption-based portfolio optimization with no credit constraints or other factors. The representative agent will allocate as a social planner would. With no frictions or externalities, markets allocate capital efficiently. This is a robust result. The pricing first order condition is already inclusive of the following, and there is an representative agent representing many investors who achieves a constrained Pareto optimal capital allocation even if the following exist:

1. The agents are unable to achieve full diversification.
2. Returns follow *any* predetermined joint pdf, including:
   (a) Returns are determined by firm-specific factors and/or loading on a systematic factor and/or multiple systematic factors.
   (b) Huge, highly correlated systematic shocks.
   (c) Individual return distributions have fat tails.
   (d) Returns are more correlated in the left-hand tails than otherwise.
3. Agents have any combination of increasing, concave utility functions.

I point-out the robustness of this pricing equation and efficient allocation by the free market to emphasize that there should be no *a priori* reason to worry about or regulate the above items. How do we reconcile this within the consumption based pricing framework? How do we, even in principle and at the most general level, differentiate between predetermined return distributions that the market will, under fairly robust assumptions, correctly price - and market failures. The distinction is what causes the mentioned return distributions.

To include ISR in returns space, I abstract from modeling particular mechanisms for ISR, and examine returns with arbitrary forms of ISR - modeled as an interfirm externality whose value depends on the stochastic performance of the firm causing the negative externality. Being careful to note the subscripts:

$$\underbrace{R_i = \underbrace{R_i'}_{\text{unimpacted returns}}}_{\text{systemically important institution}} \qquad \underbrace{R_j = \underbrace{R_j'}_{\text{unimpacted returns}} - \underbrace{f(R_i)}_{\text{systemic risk}}}_{\text{systemically impacted institution}} \tag{1}$$

Where R', the unimpacted returns, can be determined in various ways, such as with a factor model or endogenously through the representative agent's optimization. This systemic externality is very general, including any function of the returns of the systemically important institution(s). In a one-period model, $f(R_i)$ can be viewed as negatively affecting the ending cash value of the systemically impacted firm with no further interpretation needed[2]. In a multiperiod model, the ISR term $f(R_i)$ can be viewed as a causal reduction (over one time-step) to the expected sum of sdf-discounted life-time cash-flows of the firm - in other words, $f(R_i)$ is one part of the Campbell-Shiller innovation over the time-step; see Campbell and Shiller (1988).

There are several benefits to the general approach in (1), some of which are pursued by the particular manifestation in this paper. First, the implications of any model based on (1) will apply to any type of ISR. Second, because everything is in terms of returns, a consumption-based model can be solved to show how an representative agent or social planner would allocate capital to firms as part of her portfolio optimization in the presence of systemic risk. As a third benefit, deriving implications for return panels while letting the form of ISR remain general will allow the data to speak regarding what form of causal ISR is actually present.

Fourth, though the externality is in units of wealth, everything being written in terms of returns brings ISR solidly into the domain of *systematic* risk (SR). In some models, there can be difficulty cleanly distinguishing between systemic and systematic risks. There is no such problem here. This allows a factor model (see Part 3), which includes both ISR and SR, their interactions, and any implications, including return tails correlated with the industry return. Furthermore, such tails appear in the returns of the *impacted* institutions, raising questions about which firms to regulate[3].

This kind of factor model varies from most prior return specifications in the systemic risk literature (even prior return specifications that included SR and ISR) in two important ways. First, the factor model is an asset-pricing model, rather than a return-generating specification. Secondly, the factor model contains ISR as systemic risk, rather than simply adding non-Gaussian idiosyncratic or systematic terms as proxies for systemic risk.

Possibly there are parameters, γ, of the systemic risk function, $f(\cdot)$, that are impacted by characteristics of the important firm or the environment, giving $f(R_i;\gamma_j)$. Changing γ in counterfactual worlds changes the equilibria and pricing. Clearly size is one of the most important and obvious parameters. While a large number of small firms may be the source of much systemic risk in aggregate, *ceteris paribus* no single tiny firm is likely to be a large contributor. The group of firms could be modeled as the important institution or industry. A key concept of this paper is that the size of the systemic firm is an important

[2]Although information is not modeled directly here, we should view it as a full information model, where all innovations are realizations of random variables rather than discovery of information.

[3]This could be used by researchers who are developing ways to measure the contribution that a particular firm makes to overall systemic risk - such as marginal expected shortfall, derived by Acharya et al. (2010). The fact that the correlated tails might be found in the firm being affected by the systemic externality, rather than in the firm causing it, is particularly germane.

factor in determining the size of any systemic externality it may be capable of causing.

Causality necessarily brings up several issues. One is the question of what causality even means here. Aren't all returns caused by *something*? Also, if SR and ISR both cause correlation, could we ever hope to tell which is present? I discuss these and similar issues at the beginning of Part 2.

Another set of questions is about how we could model causality in returns space and distinguish it from SR. The way to model SR in returns space is well established. CAPM (Black (1972); Lintner (1965); Sharpe (1964)), the Fama-French factor model of Fama and French (1993), and the Arbitrage Pricing Theory of Ross (1976) are three examples.

The endogenous-sdf model of sections 1.2 and 1.3 let us look at the intuitions and implications of ISR for asset pricing, capital allocation, and welfare. However, analogously to SR, more structure may be needed to make certain kinds of estimates or predictions with ISR. Part 3 adds an arbitrage pricing theory factor model, based on the model presented in this paper, including ISR and SR. Factor models based on the model in this paper have many potential applications. Several are discussed. One benefit is that the distinction between ISR and SR become especially clear.

We can, perhaps more easily after looking at the factor model, characterize various finance phenomenon as being SR or ISR. Counterparty risk and margin spiral risk are clearly interfirm systemic risk. One firm holding credit default swaps written on another is as well. Financial firms levering up is unequivocal SR, not systemic risk, although it increases the probability of financial crisis and inter-industry systemic risk, as many authors have noted. Similarly, many financial firms exposing themselves to one type of risk, such as real-estate asset prices, does not represent ISR. In contrast, to the extent that there are actual economic losses, hedge funds manipulating stock prices would be an ISR. See Ben-David et al. (2013).

One result of the consumption-based model is that the existence or uniqueness of a pricing kernel when using (3) still generally follows the fundamental theorems of asset pricing (Harrison and Kreps (1979); Harrison and Pliska (1981); Delbaen and Schachermayer (1994)) regarding prices faced by the social planner. It is straight-forward to evaluate the perturbations introduced by systemic risk and their potential for disrupting the pricing kernel and equilibria (see Part 2). Though stable pricing exists for the representative agent, the fundamental theorems cannot be used to establish when. When it comes to consumption-based asset pricing, it is violation of the fundamental theorems that makes systemic risk what it is.

This Part 1 is organized as follows. The next section, 1.2, lays out the model and introduces the pricing first order conditions, first with the social planner and then with the representative agent. Theorems derive results for welfare, capital allocation, and pricing. Also with endogenous-sdf, section 1.3 discusses bailouts and how they affect the equilibrium with a real and financial industry.

The next Part (2) discusses existence and uniqueness of the pricing kernel, and of equilibrium. In the case of the representative agent, the optimum changes with her allocation, such that the fundamental theorems of asset pricing do not apply. The changing optimum also leads to a notion of stability for asset market equilibria. Part 3 puts the model into an

Arbitrage Pricing Theory factor model, using the arbitrage pricing theory for market pricing of the impacted firms. I use the model to distinguish between systematic and systemic risks, and to show how the ISR spillover decomposes naturally into additional systematic risk, additional mean-zero idiosyncratic risk, and a constant negative externality - all on the impacted firm.

1.2 Model and Pricing

1.2.1 Notation, Assumptions, Discussion of Assumptions

This is a one-period model where the investor faces a menu of investments. She will provide all funds and own the ending-period realized value of each firm, which depends on nature and the amount of funds provided to the firm. There will be a systemically important industry or institution A.

Assume utility over wealth of the standard form, $u(\cdot) \in \mathbb{C}^2$; $u'(\cdot) > 0$, $u''(\cdot) < 0$. Assume there are no borrowing or lending constraints, and full information. For this Section 1.2 and for Section 1.3, I assume all agents face complete markets (CM) and the absence of arbitrage (AOA), ensuring a unique sdf. CM and AOA are discussed more in Part 2, Section 2.2, on the existence and uniqueness of an sdf. I assume that the allocation to A, α_A, cannot exceed 1. This real asset would then take more than all of the economy's investment and have an infinite externality (outlined below). It does not occur in any equilibria, but I also assume for other situations such as comparative statics that $\alpha_A \in (0, 1)$ on the real line.

Note that the canonical portfolio optimization without ISR can have normalized initial wealth, $w_0 = 1$, without loss of generality. Note also that it has the unconstrained equivalent. The representative agent's problem with a risk-free asset with a total return of R_f, and other assets returning R_i, which are reported in excess of R_f:

$$\max_{\boldsymbol{\alpha}} E(u(\sum \alpha_i(R_i + R_f)w_0) \; s.t. \sum \alpha_i = 1 \Leftrightarrow \max_{\alpha} E[u((\sum_{i \in J} \alpha_i R_i) + R_f)]$$

$$\implies E(R_i) = \frac{-cov(R_i, u'(w_f))}{E(u'(w_f))} \tag{2}$$

where J is the set of all risky assets. R_f is the total return of the risk-free asset. All other returns are reported in excess of risk-free.

The social planner (SP) and representative agent (RA) provide the the same solution for this problem, representing efficient capital allocation by many small investors, *if* the SP is limited as outlined in Diamond's 1967 stock market paper, Diamond (1967), or if we have constant returns to scale, or returns to scale which are unknown to the SP beyond

her investment[4]. Adding systemic risk, we have some form of:

$$\underbrace{R_A = \underbrace{R'_A}_{\text{unimpacted returns}}}_{\text{systemically important institution}} \qquad \underbrace{Ri = \underbrace{R'_i}_{\text{unimpacted returns}} - \underbrace{\gamma_A\, f(R_A)}_{\text{systemic risk}}}_{\text{systemically impacted institution}} \tag{3}$$

Now add a systemically important institution A, or herd of identical institutions, that impose ISR which increases with its/their total size. Its returns in excess of risk-free are R_A and its allocation is α_A. The function $f(\cdot)$ capturing ISR is still arbitrary, but in this paper it scales up by γ_A which is increasing in the size of A. Following (3), R'_i is the unimpacted return of firm i:

$$R_i = R'_i - \gamma_A f(R_A)\ \forall i \in J\backslash\{A\},\ R_A = R'_A \tag{4}$$

where $J\backslash\{A\}$ is the set of all risky assets except for A, and $\gamma_A = {}^{\alpha_A}/_{(1-\alpha_A-\alpha_f)}$. α_f is allocation to the risk-free asset.Several of the assumptions in the model, and the basic model itself, are novel. The assumptions therefore require discussion, in the remainder of this subsection.

Regarding γ_A, a good way to capture the relative size of the systemic firm is to ask, "What is the ratio of the systemic firm's capital allocation to the the rest of the risky economy's allocation?" The question is particularly germane because ISR affects the rest of the risky economy. The exact functional form of $\gamma_A(\alpha_A)$ is not too important, but the ratio of the size of A to the rest of the risky economy makes intuitive sense and is convenient and, especially, tractable. Specifically, this ratio is $\gamma_A = {}^{\alpha_A}/_{(1-\alpha_A-\alpha_f)}$. This form has the added benefit that the ISR of A is unbounded as the firm grows to a larger portion of the economy - as the units of ISR are returns rather than currency so that a disappearing real economy could expect to receive an unbounded impact, in percentage terms. Equivalently, $\gamma_A = {}^{\alpha_A}/_{(\sum_{J\backslash\{A\}}\alpha_i)}$ is A over the rest of the risky investment set. There is no reason to think that other forms of $\gamma_A(\alpha_A)$ with $\gamma'_A(\cdot) > 0$ would yield qualitatively different results. Even so, this form being specific is a weakness of the model. Similarly, strengths would then include not using intuitive but specific forms of $u(\cdot)$, such as constant relative risk aversion or constant absolute risk aversion, or of $f(\cdot)$.

Systemically important firms add positive value as well as negative ISR. What about this value? One answer is that the firm captures all of it's added value. A few would argue that the rent-seeking activities of the financial sector imply that the firm might capture

[4]The essence of the limitation to either the SP's omniscience or omnipotence is that she cannot consider the effect of her investment on the return of the firm, so that she thinks about R_i rather than $R_i(\alpha_i) + \alpha_i * R'_i(\alpha_i)$ when optimizing. This isolates the interesting distinction, which is how the RA and SP react to ISR. The limitations in Diamond (1967) would apply to the single good in this model, equivalent to wealth. Although it works, constant returns to scale is not the best assumption here. That is because the sdf would not be truly endogenous. The sdf would only change due to the risk-free rate being endogenous. If we put the restriction instead on the SP, the model is very general. Returns could come from a production side of the economy (with any production function and industrial organization) that the agent is dealing with, or some specified returns generating process for the unimpacted returns. The only requirement is the addition of a causal externality as in (4).

more than it's added value. The assumption of capturing all value is at least reasonable. Parlour et al. (2012) assume that their financial intermediary firm captures all the value from capital transformation. They determine the value of the option to reallocate capital, and assume the intermediary captures all of this too. With this assumption, the added value only shows up in R_A.

It is reasonable that the externality is expressed as a percentage of the firm. One of several useful interpretations of the model is allocation of capital to technologies, with output as a function of input. Regardless of how big A may get, the economic value of the systemic risk externality (in say, dollars) will be limited by some notion of size for the firm being impacted. α_i is the total allocation to firm i and represents a claim to it's ending value, x. $R_j = \frac{x}{\alpha_j}$. Using (1):

$$R_j = R'_j - f(R_i) \Rightarrow x = x' - \alpha_i f(R_i)$$

where the ISR affects the ending value and is now in dollars due to scaling-up by total investment allocated to the firm.

Finally, the model is completely general to the form of $f(\cdot)$, but we might imagine that $R_i \approx R'_i$ for most realizations and $R_i < R'_i$ for critically low values of R_A. I assume only that $f(\cdot)$ is weakly positive. The ISR caused by firm or industry A does not impact A. And being risk-free, R_f is not directly affected by the stochastic ISR externality, although ISR's existence changes the risk free rate in equilibrium.

1.2.2 The Planner's Optimization

First, consider the social planner's (SP's) solution:

$$\max_{\boldsymbol{\alpha}} E(u(R_f + \sum_{i \in J} \alpha_i R'_i - \sum_{i \in J \setminus \{A\}} \alpha_i \gamma_A f(R_A)))$$

SP assumes $\gamma_A = {}^{\alpha_A}/_{(1-\alpha_A-\alpha_f)}$ is part of the optimization, so that

$$\sum_{i \in J \setminus \{A\}} \alpha_i \gamma_A f(R_A) = \sum_{i \in J \setminus \{A\}} \alpha_i \, {}^{\alpha_A}/_{(1-\alpha_A-\alpha_f)} \, f(R_A)$$

$$= {}^{(1-\alpha_A-\alpha_f)\alpha_A}/_{(1-\alpha_A-\alpha_f)} \, f(R_A) = \alpha_A f(R_A)$$

The SP's new problem is $\max_{\boldsymbol{\alpha}} E[u(R_f + (\sum_{i \in J} \alpha_i R'_i) - \alpha_A f(R_A))]$, giving the solution:

$$E(R_A) = \frac{-cov(R_A, u'(w_f))}{E(u'(w_f))} + \frac{E[f(R_A) * u'(w_f)]}{E(u'(w_f))} \tag{5}$$

$$E(R'_i) = \frac{-cov(R'_i, u'(w_f))}{E(u'(w_f))} \; \forall i \neq A \tag{6}$$

Comparative statics are a challenge. The standard method of differentiating the first order conditions, solving, and checking the Hessian is intractable.

We see in (6), the return demanded by the SP of the impacted firms depends solely on R_i', the unimpacted performance. In (5), the SP has an extra term punishing the systemically important institution for its possible ISR. The extra term is always positive, requiring *ceteris paribus* a higher expected return in order for SP to invest in the firm. Decomposing its numerator, this term is increasing in the expected value of $f(R_A)$, which is the non-size part of ISR, and goes to zero as the externality decreases and disappears. Finally, it matters how well correlated the externality is with marginal utility, also very intuitive.

Ceteris paribus above was used somewhat loosely. This is because adding ISR changes the equilibrium of this model. Normally, a fair use of *ceterus paribus* would refer to factors outside the model. Even so, the equilibrium has been perturbed in a certain way, a way where higher returns would be demanded of A for the same pdf of $u'(w_f)$ - a pdf which the existence of ISR causes us to not have. There are more rigorous comparative statics below, but it has at least as much value to think clearly about the exact perturbation that ISR causes to portfolio optimizations.

It is common for finance researchers to assume a functional form for returns that is decreasing in the amount invested in a technology, a project, or a firm - decreasing returns. Returns depend on the amount invested. Managers obviously seek to maximize returns for a given amount invested, and dislike anything that decreases returns. When the SP is investing in A, in addition to decreasing returns *per se*, we have a second decrease of expected return in the amount invested. This comes from the increase in ISR when A is allocated a larger fraction of capital.

In addition, managers prefer when markets require lower returns from their firm for a given amount invested. They prefer this if they have empire-building preferences, because the firm would be larger, and if they have compensation preferences, such that they optimally choose how much capital to accept at the terms they can get[5]. Either way, managers like lower required expected returns for a given investment. This is the interpretation on display in equation 5. ISR increases the required returns of A. This can be contrasted with the RA solution in Section 1.2.3.

The welfare effects of ISR are as we would expect:

Lemma 1. ***In the case of the SP the introduction of systemic risk is always welfare decreasing*** ***($\frac{\partial V_{SP}}{\partial c}|_{c=0} < 0$).***

Proofs for sections 1.2 and 1.3 are found in Section 1.5, starting on page 16.

The standard method of obtaining comparative statics is intractable here. Despite the addition of a positive term in the required returns of A in going from (2) to (5), the random variable $u'(w_f)$ changes from the addition of ISR, and these changes can make it possible for the SP to choose higher allocation of capital to A in the world where ISR

[5]It's probably best to focus on the empire-building preferences. One useful interpretation of the model is allocation of capital to technologies, with output as a function of input. In fact, this paper follows a common asset-pricing practice (and real-world practice) of combining corporate profits and rental rates on capital as the payoff to the owners of the firm, shareholders, who bought the firm and it's capital. The latter set of preferences, though common, would just introduce another friction, not the salient one.

is added, even though this would decrease expected returns. This may seem unlikely at first, but there are cases, somewhat akin to precautionary savings, where the SP would optimally do so. Overall increases in $u'(w_f)$, and how it covaries, make up the difference.

In fact, specific examples can be constructed that result in higher or lower allocation of capital to A by the SP after ISR is added, already a sufficient proof that the change in α_A from adding ISR cannot be signed. The representative agent's case, presented next, is less ambiguous regarding capital allocations. At first blush, we would imagine that the RA's solution would vary from the SP's only in the fact that the SP internalizes the externality while the RA fails to do so, reducing welfare. We will see below that this intuition is incorrect.

1.2.3 The Representative Agent's Optimization

In contrast to the SP, Q identical small investors solve the problem where each considers only their own contribution to γ_A from their own holding in firm A. As each investor has the same initial wealth (or the investor being considered has one Q^{th} of the wealth), we can define α'_A as the investor's proportion in A and α_A as the average proportion that the other investors have in A:

$$\gamma_A = \frac{[(Q-1)\alpha_A + \alpha'_A]}{[(Q-1)(1-\alpha_A-\alpha_F) + (1-\alpha'_A-\alpha'_F)]} \longrightarrow {}^{\alpha_A}\!/_{(1-\alpha_A-\alpha_f)} \; as \; Q \to \infty$$

The investors just take γ_A as exogenous when there are many of them. The RA solves:

$$\underset{\alpha_i \forall i \in J}{max} \; E[u(R_f + \sum_{i \in J} \alpha_i R'_i - \sum_{i \in J \setminus \{A\}} \alpha_i \gamma_A f(R_A))]$$

This provides (one form of) the first order conditions[6]:

$$E(R_A) = \frac{-cov(R_A, u'(w_f))}{E(u'(w_f))} \tag{7}$$

$$E(R'_i) = \frac{-cov(R'_i, u'(w_f))}{E(u'(w_f))} + \gamma_A \frac{E[f(R_A)u'(w_f)]}{E(u'(w_f))} \; \forall i \neq A \tag{8}$$

$$\iff E(R_i) = \frac{-cov(R_i, u'(w_f))}{E(u'(w_f))} \; \forall i \neq A \tag{9}$$

In contrast to the SP's investment in the impacted firms based upon their unimpacted performance (R'_i), the representative agent has the canonical optimization solution relative to the *impacted* performance (R_i), as shown in (9).

[6] Alternatively, as $\sum_{i \in J \setminus \{A\}} \alpha_i = (1 - \alpha_A - \alpha_F)$, the RA could solve:

$$\underset{\boldsymbol{\alpha}}{max} \; E[u(R_f + \sum_{i \in J} \alpha_i R'_i - (1-\alpha_A-\alpha_F)\gamma_A f(R_A))] \; s.t. \; \sum \alpha = 1$$

The case of no ISR (see equation 2 on page 5 if necessary) has the same equation characterizing $E(R_A)$, but does not have the additional term in (8) for $E(R_i')$, so that the set of equations here gives higher $E(R_i')$'s and a lower $E(R_A)$ than the case with no ISR, for a given $u'(w_f)$ state by state. And *ceteris paribus*, there is more capital allocated to the financial industry relative to the case of no ISR. Rigorous comparative statics show that this is not just *ceteris paribus*, but always. This is not simply relative to the case of the SP. This is relative to the case of the case of no ISR. The market does not ignore the externality. The market, surprisingly, exacerbates the problem. Comparative static:

Theorem 1. ***The introduction of systemic risk will increase the RA's allocation to A, increasing systemic risk further*** ($\frac{\partial \alpha_{A,RA}}{\partial c}|_{c\in\varepsilon(0+)} > 0 \; \forall \alpha_A < 1$)**.**

where we use $c \in \varepsilon(0+)$, a neighborhood around $c = 0$ with $c > 0$, due to the first derivative being zero at $c = 0$. Comparison with the SP, who internalizes the externality, is worse. To review the SP: She took account of the externality and assigned it to the systemically impacting firm. The demand by the SP for higher returns *ceteris paribus* from A appears in the positive term added to the returns of A in (5). The punishment scales with the magnitude of $f(\cdot)$ and with how well it covaries with marginal utility.

Rather than simply the removal of this positive term in the equation for $E(R_A)$ by the RA, as if the market were failing to take the externality into account, the market ignores the externality for A. The RA then makes the problem worse by attributing it to the impacted firms[7]. Whether they have compensation or empire-building preferences, the managers of A have no incentive to reduce the ISR of their firm.

The effects on welfare can now be introduced. The SP always does better than an agent who ignores the externality, and such an agent does better than the RA:

Corollary 1. ***In the case of the RA, the introduction of systemic risk is always welfare decreasing*** ($\frac{\partial V_{RA}}{\partial c}|_{c=0} < 0$)**.**

This was true for the SP as well.

[7] The alternative maximization mentioned in the prior footnote can be solved with reasonable assumptions on the bond market to give:

$$E(R_A) = \frac{-cov(R_A, u'(w_f))}{E(u'(w_f))} - \gamma_A \frac{E[f(R_A)u'(w_f)]}{E(u'(w_f))}$$

$$E(R_i') = \frac{-cov(R_i', u'(w_f))}{E(u'(w_f))} \; \forall i \neq A$$

These equivalent first order conditions are marginally easier to compare with those of the SP. Here, the market, as many investors, simply ignores the externality when it comes to the firms i. Something more subtle and unexpected happens with the impacting firm however. Its expected return equation has a term beyond the standard solution with externality ignored. There is an extra term that can be signed as negative in the equation for $E(R_A)$. In these first order conditions, rather than removal of the positive term in $E(R_A)$, a negative one takes its place, more directly rewarding the systemically risky firm for causing degradation of the investor's other options.

Corollary 2. ***The introduction of systemic risk is less welfare decreasing for the SP than for the RA*** ***(*** $\frac{\partial V_{RA}}{\partial c}|_{c=0} < \frac{\partial V_{SP}}{\partial c}|_{c=0}$ ***).***

This corollary would be true even for a situation that an SP internalizes an externality and the RA does not. We know from theorem 1 that the situation is worse than that.

When there is an unpriced externality, normally government intervention can, at least in theory and sometimes in practice, charge someone for the ignored externality and return to an efficient outcome. In light of the market not just ignoring this externality, it is not immediately clear what the results of a simple charge or capital charge will be.

The welfare ordering from highest to lowest is slightly different than with an externality model where the market ignores an externality:

1. The case of no ISR, which is the same for RA and SP.

2. SP when ISR is present.

3. When ISR is present, the welfare that results from RA not knowing ISR is present, and allocating as if it were not.

4. The case of ISR for RA, corresponding to the situation we usually have, giving the least efficient of these allocations.

In the next section, this situation presents a challenge to the government when bailouts are used to mitigate the systemic damage. This is not the well-known challenge of dealing with firms' decisions and the resulting moral hazard (managers taking on excessive risk with the belief that the government will have to bail them out if they fail).

1.3 Mitigating Systemic Risk with Bailouts

1.3.1 The General Case

Assume that the realized ISR can be mitigated by bailing out the offending firm *ex post*. The bailout amount is defined as $\alpha_A \Lambda$, and this cost is reduced from the RA's ending wealth. The assumption is that the government takes this money from the RA in cash and spends it on mitigation of the ISR externality. Bailout is ex post optimally determined by:

$$\max_{\Lambda} \; u(R_f + \sum_{i \in J} \alpha_i R_i' - \underbrace{\alpha_A \Lambda}_{\text{cost}} - \sum_{i \in J \setminus \{A\}} \alpha_i \gamma_A f(R_A + \Lambda))$$

The bailout has increased the argument of $f(\cdot)$ as the bailout amount, $\alpha_A \Lambda$, could increase the returns *ex post* of firm A by Λ. However, the returns of firm A have not been increased within the term $\sum \alpha_i R_i'$, capturing the idea that the firm is not benefiting from the bailout in order to avoid moral hazard problems with how much risk the firm takes; R_A itself is not changing, but the government's expenditure makes ISR goes from $\gamma_A f(R_A)$ down to $\gamma_A f(R_A + \Lambda)$. Recalling $\sum_{i \in J \setminus \{A\}} \alpha_i \gamma_A = \alpha_A$, ex-post optimal bailout, Λ^*, occurs when $f'(R_A + \Lambda^*) = -1$, with the restriction that $\Lambda^* \geqq 0$. As mentioned, it is reasonable to assume that $f(\cdot)$ has $f(\cdot) \geqq 0$, $f'(\cdot) \leqq 0$, $f''(\cdot) \geqq 0$, and $f(\cdot) = 0$ for positive arguments. In this case, there would be a unique solution either fulfilling the first order condition, or at $\Lambda^* = 0$ if $f'(R_A) > -1$. $\Lambda^* = max\{0,\ f'^{-1}(-1) - R_A\}$. The results in this section also assume $f'(R_A)$ has positive support below -1.

If optimal *ex post* bailout is anticipated by the RA, then the allocation problem for the RA is replaced by:

$$\underset{\boldsymbol{\alpha}}{max}\ E[u(R_f - \alpha'_A\Lambda^* + \sum_{i\in J}\alpha_i R'_i - \sum_{i\in J\setminus\{A\}}\alpha_i\gamma_A f(R_A+\Lambda^*))]$$

Even though the government gets the money for the bailout from the RA after returns have been realized, the RA does not consider the cost of the bailout ($\alpha'_A\Lambda^*$) in her optimization because each agent considers her contribution to be insignificant. Writing α'_A means that it is treated as exogenous by the RA. The first order conditions are:

$$E(R_A) = \frac{-cov(R_A, u'(w_f))}{E(u'(w_f))} \tag{10}$$

$$E(R'_i) = \frac{-cov(R'_i, u'(w_f))}{E(u'(w_f))} + \gamma_A\frac{E[f(R_A+\Lambda^*)u'(w_f)]}{E(u'(w_f))}\ \forall i \neq A \tag{11}$$

$$\iff E(\tilde{R}_i) = \frac{-cov(\tilde{R}_i, u'(w_f))}{E(u'(w_f))}\ \forall i \neq A \tag{12}$$

where $\tilde{R}_i = R'_i - \gamma_A f(R_A+\Lambda^*)$ is the return of asset i as the RA sees it, knowing about the ISR and the bailout rule. When comparing it to the solution without bailouts, the second term in (11) ($\gamma_A\frac{E[f(R_A+\Lambda^*)u'(w_f)]}{E(u'(w_f))}$) is identical except that it includes Λ^*, and is also positive (see equation 10 if necessary). Because $f(\cdot)$ is decreasing, the second term in (11) for the required return $E(R'_i)$ is smaller than when no bailout is anticipated.

In the case above, the only thing that had changed for the SP's optimization was the expected ISR. Now we consider the case that shareholders are allowed to benefit. This means that the returns of firm A have been increased within the term $\sum\alpha_i R'_i$.

$$\underset{\boldsymbol{\alpha}}{max}\ E\{u[R_f - \alpha'_A\Lambda^* + \alpha_A(R_A+\Lambda^*) + \sum_{i\in J\setminus\{A\}}(\alpha_i R'_i - \alpha_i\gamma_A f(R_A+\Lambda^*))]\}$$

$$E(R_A) = \frac{-cov(R_A, u'(w_f))}{E(u'(w_f))} - \frac{E[\Lambda^* u'(w_f)]}{E(u'(w_f))} \tag{13}$$

$$E(\tilde{R}_i) = \frac{-cov(\tilde{R}_i, u'(w_f))}{E(u'(w_f))}\ \forall i \neq A \tag{14}$$

The first order condition for the R_i's did not change. Note the subtraction from $E(R_A)$ which is not present when firm A is not allowed to benefit from the bailout. This lowers the required returns of the firm, *ceteris paribus*.

Returning for the remainder of this section to the case where the firm does not benefit, it is not surprising that the SP always benefits from bailout:

Lemma 2. ***Introducing bailout increases welfare for the*** SP ***(*** $\frac{\partial V_{SP}}{\partial c}|_{c=0} > 0$ ***)***

The addition of bailout is represented by replacing Λ^* with $c\Lambda^*$. The increase in welfare is:

$$\frac{\partial V_{SP}}{\partial c}|_{c=0} = -E[u'(w_f)\alpha_A^*(c)(f'(R_A)+1)\Lambda^*]$$

where α_A^* is a function of c. And the bailout, Λ^*, was defined as $\Lambda^* = max\{0,\ f'^{-1}(-1) - R_A\}$.

Lemma 3. ***If adding bailout decreases the*** RA***'s allocation to*** A***, then it increases welfare*** ($\frac{\partial \alpha_A^*}{\partial c}|_{c=0} < 0 \Rightarrow \frac{\partial V_{RA}}{\partial c}|_{c=0} > 0$)

Economically, the RA is overallocating to A, so if bailout alleviates this ($\frac{\partial \alpha_A^*}{\partial c}_{c=0} > 0$), then welfare is improved.

For the RA, dropping the arguments on $\alpha_A^*(c)$, $\gamma_A(c)$, $f(R_A)$ and $f'(R_A)$:

$$\frac{\partial V_{RA}}{\partial c}|_{c=0} = \underbrace{-E[u'(w_f)\alpha_A^*(f'+1)\Lambda^*]}_{\text{direct effect}} + \underbrace{-E\{u'(w_f)(1+\gamma_A)\, f\, \frac{\partial \alpha_A^*}{\partial c}|_{c=0}}_{\text{distortion effect}}$$

Corollary 3. ***The direct effect is always positive*** ($-E[u'(w_f)\alpha_A^*(f'+1)\Lambda^*] > 0$)

Also, the sign of the distortion effect depends only on the sign of $\frac{\partial \alpha_A^*}{\partial c}|_{c=0}$, having the opposite sign. If bailout exacerbates over-allocating to A, ($\frac{\partial \alpha_A^*}{\partial c}|_{c=0} > 0$), then the distortion effect lowers welfare. If the distortion effect is great enough, then $\frac{\partial V_{RA}}{\partial c}|_{c=0} < 0$. One can find specific examples of both signs. The sign of $\frac{\partial V_{RA}}{\partial c}|_{c=0}$ is ambiguous. Because the assumptions to this point are broad, a claim of sign ambiguity is not a strong statement. It could be only very obscure cases with one sign or the other.

1.3.2 Two-industry Model

In this subsection, I continue with the case that the one receiving the bailout is not allowed to benefit. I consider the case with quadratic utility, a systemically risky industry (the financial industry), and a single impacted industry (the real industry) to gain further intuition about financial-industry bailouts. Also, the risk-free rate will continue to be endogenous, but some specification of the bond market is required: For this section, I assume that the risk-free asset, being a purely financial asset, is in zero net supply. There is no risk-free technology.

Theorem 2. There are cases where adding bailout decreases welfare for the RA**.** ($\frac{\partial V_{RA}}{\partial c}|_{c=0} < 0$)

As described above this comes from cases where allocation to A increases and the distortion effect is of greater magnitude than the direct effect. This says that *ex post* optimal bailouts to the financial industry can distort capital allocation to the point that the distortion increases the damage of a financial crisis more than bailout decreases the damage of a financial crisis, in expectation. This may initially be a surprise when we consider that the firm's investors do not benefit from the bailout.

Corollary 4. Adding bailout decreases welfare iff: $0 < E(R_A^2 - R_A R_B' - R_A f\,) < \kappa$
where $\kappa \equiv \dfrac{E[u'(w_f)(1+\gamma_A)\,f]\;E[\alpha_A R_A \Lambda^*(f'+1)]}{-E[u'(w_f)\alpha_A(f'+1)\Lambda^*]} \quad (>0)$

We would image $E(R_A R_B') > 0$. So

$$[\underbrace{E(R_A^2)}_{>0} - \underbrace{E(R_A R_B')}_{>0} - \underbrace{E(R_A f)}_{<0}\,]$$

being within some positive range spanning from zero is clearly plausible. It seems more likely that this expression is positive than negative: There is the positive third term, and there is the probability that the squared returns of the financial industry are at least as high as the product of its returns with the unimpacted returns of the real industry. One likely scenario that it is positive, and high enough, so that the expression is greater than κ and bail-outs are effective. Another is that it is positive, but too low, so that bail-outs are counter-productive.

In such a situation, bail-outs are effective when systemic risk is large and covaries with the financial industry, when the financial industry has high volatility, high return, and/or low correlation with the real industry. If the volatility of the financial industry were low, this difference could even be negative, again making bailouts effective. It is worth emphasizing that the owners of the systemic industry do not benefit from the bail-out and the mechanism making the bail-out ineffective is that it increases capital allocation distortion. The distortion effect is just the change in distortion. The total distortion from ISR could be large.

Also, it is not the usual moral hazard argument, as firms do not have choices here. There is no endogenous risk-taking in this set-up. The core intuitions are about incentives in the RA's portfolio problem, not firm behavior. Endogenous firm risk-taking could be added to this framework, but my preliminary examination appears to show that there's no special intuition around the interaction between the manager's moral hazard and the RA optimization. The moral hazard around risk-taking has been well covered in simpler models. There is more generality to these several bailout intuitions than may appear at first blush. The models could be changed to accommodate an arbitrary level of mitigation and an arbitrary level of benefit to the bailed-out firm, but the added complexity yields no new intuition. This would be accomplished by replacing $R_A + \Lambda$ with $R_A + g(\Lambda)$, where $g(\cdot)$ is increasing and captures the amount that the firm is allowed to benefit. The next step is to replace $f(R_A + \Lambda)$ with $f(R_A + h(\Lambda))$, where $h(\cdot)$ is increasing and captures the amount that the bailout mitigates externalities. We could even imagine, at least in theory, a case where $h(\Lambda) > \Lambda$, such as when the government gives money to the firm but requires the money to be spent in ways that are particularly effective in reducing ex-post ISR. With these two changes, the cost of the bailout, $\alpha_A \Lambda$, would not change.

This set-up accommodates any mitigation and benefit possibilities the government could face. It can also accommodate program costs or other deadweight losses in getting the money to the firm or extracting the money from the RA, presumably via taxation.

The solutions are only slightly more complicated but qualitatively identical to what has been presented[8].

1.4 Part 1 Conclusion

This Part contributed a model of the canonical portfolio optimization problem with the addition of interfirm systemic risk added. The addition was a general functional form of returns yielding the externality, also as returns. The generality can accommodate any form of interfirm systemic risk (ISR) that can be written in terms of returns. In order to assess issues of capital allocation, asset prices, investment and welfare; the particular manifestation of the model considered in this paper had a systemically important institution with an externality that increased as firm size increased.

The model was considered with a social planner (SP) or representative agent (RA) performing the optimization. The base-case results are without ISR: if there is no externality (or equivalently if it is being ignored), the SP and RA reach the same optimal allocation. A key result of the case with ISR is that the investment by the RA in the systemically risky firm is not just more than if the RA were appropriately accounting for the systemic externality, it is more even than if the RA were ignoring the externality. The investment of the RA exacerbates the market failure rather than just ignoring it, and this effect increases in the amount of ISR. The intuition for the result is that the threat of ISR onto the real economy makes outside options worse and increases investment in, and the size of, the financial industry or systemically important firm. The results would appear to show that the welfare loss from the capital allocation distortion that systemic risk causes, is significant[9]. The portfolio optimization model with endogenous-sdf in this Part 1 uncovered many new financial and economic intuitions about asset prices, capital allocation, welfare, and systemic risk - as well as bail-outs and their potential for direct effects and allocation distortion effects.

There are many possible extensions to the model. The most useful is probably a model with multiple systemically important firms. Perhaps such a model could be a first step toward a consumption-based optimization of networks. Network models, such as Allen and Gale (2000); Gale and Kariv (2007); and Gai and Kapadia (2010), generally use simple modal (0,1) utility functions. One factor in intermediation is the funds that the intermediary has available. This can be reasonably modeled as a function of the returns of the intermediary. A negative shock to returns decreases firm funds and the ability to borrow, causing a negative externality. Another possibility is a model with externalities as a function of the returns of many firms, affecting the returns of many firms.

[8]The cases that were presented above correspond to $g(\Lambda) = 0$, $h(\Lambda) = \Lambda$ for the first case where the firm is not allowed to benefit, and $g(\Lambda) = h(\Lambda) = \Lambda$ for the second case where the firm benefits directly.

[9]An announced *ex post* optimal bail-out has a beneficial direct effect. It also changes the level of capital allocation distortion from ISR, an effect which can be beneficial or detrimental. The bail-out's worsening of the capital allocation distortion can be so great that the bail-out is counter-productive. The conditions under which this can happen are surprisingly reasonable and possible in some economies, while this is just the *change* in the distortion effect.

Another potential application of an endogenous-sdf model with many systemically important firms is modeling the financial industry. The model in this paper applies fairly directly to the case of the financial industry imposing a possible negative externality on the other industries. What's missing is the internal workings of the financial industry. Using a model in a model, the industry could be modeled with multiple systemically important firms (with the endogenous-sdf model used in the previous sections) to give a pdf of the returns of the financial industry (net of all SR and ISR)[10]. With an assumed systemic risk function, $f(\cdot)$, the set-up can be imported into the consumption-based model, with two industries, or imported into the simpler two-industry model from section 1.3.2. This may be doable, or it may require simplification (such as the factor model in Part 3) for tractability.

1.5 Part 1 Proofs

Section 1.2:

Lemma 1: In the case of the *SP* the introduction of systemic risk is always welfare decreasing ($\frac{\partial V_{SP}}{\partial c}|_{c=0} < 0$):

Proof:

$$\max_{\alpha} \; E(u(R_f + \sum_{i \in J} \alpha_i R_i' - \sum_{i \in J \setminus \{A\}} \alpha_i c \gamma_A f(R_A)))$$

SP assumes γ_A is part of the optimization, so that:

$$\sum_{i \in J \setminus \{A\}} \alpha_i = 1 - \alpha_A - \alpha_f \Rightarrow$$

$$\sum_{i \in J \setminus \{A\}} \alpha_i \gamma_A c f(R_A) = (1 - \alpha_A - \alpha_f)^{\alpha_A / (1 - \alpha_A - \alpha_f)} c f(R_A) = \alpha_A c f(R_A) \tag{15}$$

$$V_{SP}(c) \equiv \max_{\alpha} \; E(u(R_f + (\sum_{i \in J} \alpha_i R_i') - \alpha_A c f(R_A))) \Rightarrow$$

$$V_{SP}(c) = E(u(R_f + (\sum_{i \in J} \alpha_i^* R_i') - \alpha_A^* c f(R_A))) \tag{16}$$

$$\Rightarrow \frac{\partial V_{SP}}{\partial c}|_{c=0} = -E[u'(w_f) \alpha_A f(R_A)] < 0$$

by the envelope theorem.

Section 1.2.2 noted that the SP's allocation to A can increase, decrease, or stay the same with the introduction of ISR. Next I show that introduction of the ISR externality

[10] When put into the framework of this paper, risk sharing between the firms in the industry should be considered ISR. This is because the firms are not just loading up on one particular risk factor, they are contracting with each other to share risks. Thus, at least one paper has considered the impact of one type of ISR, as well as SR, on the probability of a financial crisis Ibragimov et al. (2011). The authors did not use a concave utility function though, but did other things.

will always cause the RA to increase investment in the firm causing the externality, rather than just failing to take it into account by keeping her same portfolio:

Theorem 1: The introduction of systemic risk will increase the RA's allocation to A ($\frac{\partial \alpha_{A,RA}}{\partial c}|_{c\in\varepsilon(0+)} > 0 \; \forall \alpha_A < 1$):

First note that $\sum_{i\in J}$ is over all risky assets, including A, and $\sum_{i\in J\setminus\{A\}}$ is over all the risky assets except A, i.e. it's over all the risky assets whose returns are reduced by: $-\gamma_A f(R_A)$. So, $\sum_{i\in J}\alpha_i = 1-\alpha_F$, and $\sum_{i\in J\setminus\{A\}}\alpha_i = 1-\alpha_A-\alpha_F$. Also, $\gamma_A \equiv {}^{\alpha_A}/_{(1-\alpha_A-\alpha_f)}$.

$$V_{RA}(c) \equiv \max_{\alpha} \; E(u(R_f + \sum_{i\in J}\alpha_i R_i' - \sum_{i\in J\setminus\{A\}}\alpha_i c\gamma_A f(R_A)))$$

$$= E(u(R_f + \sum_{i\in J}\alpha_i^* R_i' - \sum_{i\in J\setminus\{A\}}\alpha_i^* c\gamma_A f(R_A)))$$

For finding the partial, we work with the following expression for $V_{RA}(c)$, which separates terms that are endogenous and exogenous to the optimizing agent. Employing the envelope theorem, only for terms that have been maximized over (not γ_A):

$$\frac{\partial V_{RA}}{\partial c} = -E[u'(w_f)(\alpha_A f(R_A) + (1-\alpha_A-\alpha_F)\frac{\partial \gamma_A}{\partial c} c f(R_A))]$$

$$(\; where \;\; \frac{\partial \gamma_A}{\partial c} = [\frac{1}{(1-\alpha_A)} + \frac{\alpha_A}{(1-\alpha_A)^2}]\frac{\partial \alpha_A}{\partial c} \Rightarrow (\; \frac{\partial \gamma_A}{\partial c} > 0 \Leftrightarrow \frac{\partial \alpha_A}{\partial c} > 0 \;)\;)$$

where we used (15). Although it would not help with the derivative, we could also write $V_{RA}(c)$ as:

$$V_{RA}(c) = E(u(R_f + (\sum_{i\in J}\alpha_i^* R_i') - \alpha_A^* c f(R_A))) \tag{17}$$

The expression for $V_{RA}(c)$ in (17) is the same as that for $V_{SP}(c)$ in (16), but generally $V_{RA}(c) \neq V_{SP}(c)$ because the α^*'s come from a different first order condition. The standard portfolio allocation solution without ISR (equation 2 on page 5) is the solution for both the SP and the RA, so that $\boldsymbol{\alpha^*_{SP}} = \boldsymbol{\alpha^*_{RA}}$ when $c = 0$.

$$V_{RA}(0) = V_{SP}(0) = E(u(\sum_{i\in J}\alpha_i^* R_i' - \alpha_A^* 0 f(R_A) + R_f))$$

As c is increased from 0, their portfolio allocations will differ, as they have different first order conditions, equations (5) - (9). Now we examine what happens as we approach $c = 0$ from the right, $c = \triangle c$. For $c > 0$, no matter how small, $V_{SP}(c)$ is the maximum value of $E(u(\sum_{i\in J}\alpha_i R_i' - \alpha_A c f(R_A) + R_f))$ that can be achieved, due to the definition of $V_{SP}(c)$ as the maximization of this very thing. Hence:

$$\forall c \neq 0, V_{SP}(c) > V_{RA}(c) \Rightarrow \forall c \neq 0, V_{SP}(c) - V_{RA}(c) > 0; \;\; V_{SP}(0) = V_{RA}(0)$$

$$\Rightarrow \frac{V_{SP}(\triangle c) - V_{RA}(\triangle c)}{\triangle c} > 0 \Rightarrow \frac{[V_{SP}(\triangle c) - V_{SP}(0)] - [V_{RA}(\triangle c) - V_{RA}(0)]}{\triangle c - 0} > 0$$

$$\Rightarrow lim\ \triangle c \to 0^+ \frac{[V_{SP}(\triangle c) - V_{SP}(0)] - [V_{RA}(\triangle c) - V(0)]}{\triangle c} > 0$$

And because the $V(\cdot)$'s are the max functions of well-behaved problems, $V(\cdot) \in \mathbb{C}$, $\mathbb{C}^1$:

$$lim \triangle c \to 0^+ \frac{[V_{SP}(\triangle c) - V_{SP}(0)] - [V_{RA}(\triangle c) - V(0)]}{\triangle c} > 0 \iff \frac{\partial V_{SP}}{\partial c}|_{c=0^+} - \frac{\partial V_{RA}}{\partial c}|_{c=0^+} > 0$$

$$\Rightarrow \frac{\partial(\Delta V)}{\partial c} = -E[u'(w_f)(\alpha_A f(R_A)] + E[u'(w_f)(\alpha_A f(R_A) + (1 - \alpha_A - \alpha_F)\frac{\partial \gamma_A}{\partial c} c f(R_A))]$$

$$\frac{\partial(\Delta V)}{\partial c}|_{c \in \varepsilon(0+)} > 0 \Rightarrow \frac{\partial \alpha_A}{\partial c}|_{c \in \varepsilon(0+)} > 0\ \forall \alpha_A < 1$$

$\frac{\partial \alpha_A}{\partial c}|_{c=0} < 0\ \forall \alpha_A > 1$, is just an artifact of the model where γ_A switches sign above $\alpha_A = 1$. By the time α_A reaches one, γ_A has, by design, diverged to infinity and the economy is long past financial crisis. This is why I defined $\alpha_A \in (0, 1)$.

Corollary 1: In the case of the *RA*, the introduction of systemic risk is always welfare decreasing ($\frac{\partial V_{RA}}{\partial c}|_{c=0} < 0$)

As above.

Corollary 2: The introduction of systemic risk is less welfare decreasing for the *SP* than for the *RA* ($\frac{\partial V_{RA}}{\partial c}|_{c=0} < \frac{\partial V_{SP}}{\partial c}|_{c=0}$)

As above.

Section 1.3:

Lemma 2: Adding bailout increases welfare for the SP ($\frac{\partial V_{SP}(c)}{\partial c}|_{c=0} > 0$)

$$V_{SP}(c) = E[u_{SP}(R_f + \sum_{i \in J} \alpha_i^* R_i' - \alpha_A^* c\Lambda^* - \alpha_A^* f(R_A + c\Lambda^*)],\ \Lambda^* = max\{0,\ f'^{-1}(-1) - R_A\}$$

$$\frac{\partial V_{SP}}{\partial c}|_{c=0} = -E[u'(w_f)\alpha_A^*(c)(f'(R_A) + 1)\Lambda^*]$$

where $\alpha_A^*(c)$ is a function of c. Each element inside the expectation, including $[-\Lambda^*(f'(R_A + \Lambda^* c) + 1)]$, is weakly positive for all states, ω. If there is no bailout, $\Lambda^* = 0$. If there is bailout, $f'(R_A + \Lambda^*) = -1$ and $f'(R_A + c\Lambda^*) < -1\ \forall c \in [0, 1)$, so the product $[\Lambda^*(f'(R_A) + 1)] \leqq 0$, $\forall \omega$.

Lemma 3: If adding bailout decreases the RA's allocation to A, then it increases welfare ($\frac{\partial \alpha_A^*}{\partial c}|_{c=0} < 0 \Rightarrow \frac{\partial V_{RA}}{\partial c}|_{c=0} > 0$)

For the RA, dropping the arguments on $\alpha_A^*(c)$ and $\gamma_A(c)$, $f(R_A)$ and $f'(R_A)$:

$$\frac{\partial V_{RA}}{\partial c}|_{c=0} = \underbrace{-E[u'(w_f)\alpha_A^*(f'+1)\Lambda^*]}_{\text{direct effect}} + \underbrace{-E\{u'(w_f)(1+\gamma_A)\, f\, \frac{\partial \alpha_A^*}{\partial c}|_{c=0}}_{\text{distortion effect}}$$

note that $-u'(w_f)(1+\gamma_A)\, f$ is negative for all states.

Corollary 3: The direct effect is always positive ($-E(u'(w_f)\alpha_A^*(f'+1)\Lambda^*) > 0 \forall c = 0$)

As above in Lemma 2.

Theorem 2: There are cases where adding bailout decreases welfare for the RA. ($\frac{\partial V_{RA}}{\partial c}|_{c=0} < 0$)

$$\frac{\partial V_{RA}}{\partial c}|_{c=0} = -E[u'(w_f)\alpha_A^*(f'+1)\Lambda^*] + -E\{u'(w_f)(1+\gamma_A)\, f\, \frac{\partial \alpha_A^*}{\partial c}|_{c=0} \tag{18}$$

Dropping the the subscript from $\frac{\partial \alpha^*}{\partial c}|_{c=0}$, consider the first order condition, $E[u'(w_f)R_A] = 0$:

$$\frac{\partial E[u'(w_f)R_A]}{\partial c}|_{c=0} = -E\{u''(w_f)R_A[(\alpha_A^*\Lambda^* - \sum_{i\in J}\frac{\partial \alpha_i^*}{\partial c}R_i' + \frac{\partial \alpha_A^*}{\partial c}f + \alpha_A^* f'\Lambda^*]\} = 0$$

$$= -E\{u''(w_f)R_A[(f - R_A)\frac{\partial \alpha_A^*}{\partial c} + \alpha_A^*\Lambda^* - \sum_{i\in J\setminus\{A\}}\frac{\partial \alpha_i^*}{\partial c}R_i' + \alpha_A f'\Lambda^*]\} = 0$$

Now add the restrictions of quadratic utility, a systemically important industry, and a systemically impacted industry, B:

$$-E\{u''(w_f)R_A[(f - R_A)\frac{\partial \alpha_A^*}{\partial c} - \frac{\partial \alpha_B^*}{\partial c}R_B' + \alpha_A\Lambda^*(f'+1)]\} = 0$$

$$-E(R_A(f - R_A + R_B'))\,\frac{\partial \alpha_A^*}{\partial c} - E(R_A\alpha_A\Lambda^*(f'+1)) = 0$$

$$\frac{\partial \alpha_A^*}{\partial c} = \frac{E(\alpha_A R_A\Lambda^*(f'+1))}{E(R_A^2 - R_A R_B' - R_A f)} \tag{19}$$

here we used that $\frac{\partial \alpha_A^*}{\partial c} = -\frac{\partial \alpha_B^*}{\partial c}$ and $u''(w_f)$ is constant. Recall $f(\cdot)$ has $f(\cdot) \geqq 0$, $f'(\cdot) \leqq 0$, $f''(\cdot) \geqq 0$, and $f(\cdot) = 0$ for positive arguments. So, we have states with bailout: R_A, $(f'+1) < 0$; f, $\Lambda^* > 0$; states without bailout where R_A is positive: $\Lambda^* = f = 0$; and

states without bailout where R_A is negative: $\Lambda^* = 0,\ f \gtreqless 0$. The first of the following inequalities are in (18) and the second are in (19):

$$-E[u'(w_f)\alpha_A^*(f'+1)\Lambda^*] > 0,\ \ E[u'(w_f)(1+\gamma_A)\,f)] > 0$$

$$E(\alpha_A R_A \Lambda^*(f'+1)) > 0,\ \ E(R_A f)\ < 0$$

Combining (18) and (19):

$$\frac{\partial V_{RA}}{\partial c}|_{c=0} < 0 \Leftrightarrow$$

$$\frac{1}{E(R_A^2\ - R_A R_B'\ - R_A f\,)} > \frac{-E[u'(w_f)\alpha_A^*(f'+1)\Lambda^*]}{E[u'(w_f)(1+\gamma_A)\,f]\ E(\alpha_A R_A \Lambda^*(f'+1))}\ (>0\,)$$

$$0 < E(R_A^2\ - R_A R_B'\ - R_A f\,) < \frac{E[u'(w_f)(1+\gamma_A)\,f]\ E(\alpha_A R_A \Lambda^*(f'+1))}{-E[u'(w_f)\alpha_A^*(f'+1)\Lambda^*]} \equiv \kappa\ (>0\,)$$

Corollary 4: Adding bailout decreases welfare iff: $0 < E(R_A^2 - R_A R_B' - R_A f\,) < \kappa$

As above.

2 Equilibrium

2.1 Epistemology

What type of equilibrium is being characterized? There are several epistemological issues here. They regard what causality means in this context and how the above model really differs from any other specification of the prospects of a firm, specified by cash-flows through time and across states of nature - cash-flows which are presumably always caused by something. Consider the following line of questioning: The distributions of the cash-flow prospects of firms arise for a wide variety of reasons. Even a completely exogenous model of cash-flows presumes that those cash-flows are caused.

Three common questions arise: What does it even mean to say that there is a causal externality? How could we ever hope to tell? What possible implications could it have for anyone's behavior? To discuss the issues at hand and see the distinctions, we'll introduce a toy model with a simple and laid-out return-generating process. In the end, none of the concepts will be specific to this simple model.

To get our finger completely on this issue, consider the following two unlikely models of an economy consisting of two firms. Assume that there is an exogenous macro-factor, λ, that is very important to the performance of both firms. In fact, in the absence of ISR returns are determined exclusively from this factor and an idiosyncratic shock. Now we add an interfirm systemic externality into the returns of the systemically impacted firm j, which is a function of the returns of the systemically important institution i. This can be done in two ways, generating the following two economies:

$$R_i = \zeta_i \lambda + \varepsilon_i \qquad\qquad R_j = \zeta_j \lambda - f(R_i) + \varepsilon_j$$

or:

$$R_i = \zeta_i \lambda + \varepsilon_i \qquad\qquad R_j = \zeta_j \lambda - f(\zeta_i \lambda + \varepsilon_i) + \varepsilon_j$$

Using this simple model, we can ponder those common epistemological questions that arise:

1. What does it even mean to say that there is a causal externality? This now has a clear answer. *It means that in a counterfactual world where we exogenously modify only R_i, the returns of R_j would be different.* This is almost never true of a portfolio optimization problem. Does the agent, even the SP, ever consider what her investment in one company could do to her investment in another?

This would not be true in the second model; an exogenous alteration in a counterfactual world is **not** part of ε_i. This is the logic behind a bailout for instance, to mitigate the effects. In the second model, we just have a joint pdf specification and no basis for a bailout. The pdf of R_j has changed only in the former case.

2. How could we ever hope to tell? One way is to observe the causation in action. Another might be to look for cases of exogenous intervention into the returns of firms that we think are causing externalities and try to isolate the effects of the interventions. A third way is that, unlike in our toy model, causal connections should be expected to have different joint pdf's. If we go back to the most general model:

$$\underbrace{R_i = \underbrace{R_i'}_{\text{unimpacted returns}}}_{\text{systemically important institution}} \qquad\qquad \underbrace{R_j = \underbrace{R_j'}_{\text{unimpacted returns}} - \underbrace{f(R_i)}_{\text{systemic risk}}}_{\text{systemically impacted institution}}$$

We could assume that returns of the firms come from idiosyncratic risk, systematic risk, and systemic risk. The factor causing the systemic risk is the sum of many small factors and is likely to be close to normal Casella and Berger (2001). The same is true of idiosyncratic risk. In the absence of ISR, we should expect the returns to be normal. But ISR comes from a transformation of a normal. And most specific models of ISR would imply that the transformation is strictly increasing, weakly negative, strictly concave, close to zero for most return realizations, and largely negative for very low return realizations. This would cause fat left-hand tails in the impacted firm and correlated left-hand tails between the firms, matching observation, especially in the financial industry.

3. The third question of would there be any implications. The question is: if this is true, do we do something different? And the answer here is a resounding yes. If these are connections are causal, then there are welfare gains to limiting them with regulation or directing capital more efficiently. In addition, we can see in the above that the externality has two problems. The first is that it is possibly a market failure itself. The second is that, along with SR, it is counter-cyclical and increases the chance of a large negative return realization in the financial industry as a whole, possibly causing a financial crisis and negative externality to the real economy.

It would also imply that proposals to regulate firms that have fat and correlated left-hand tails may be regulating the victim. What if we were allowed to regulate some firms from the portfolio comprising the financial industry in order to decrease the chance of a

very negative return realization of the industry as a whole? It is true that if the pdf's of returns were independent, we should get focus on firms with fat and correlated left-hand tails as they have poor performance at the wrong time. But the model just discussed has larger implications about who should be regulated.

2.2 Existence and Uniqueness

2.2.1 Set-up

To examine what ISR might do to the stability of asset markets and equilibria, the best way is to begin with a group of assets that would yield a unique stochastic discount factor (sdf) in the absence of ISR, and then ask what adding ISR does to existence and uniqueness. Any group of assets that would yield a unique sdf in the *absence* of ISR will be represented by an S matrix. Another way to think of S is as a matrix of unimpacted returns. The group of assets faced by the SP or RA after adding ISR to an S matrix will be represented by a T matrix.

We will then be investigating existence and uniqueness of equilibria for the T matrix, in order to ask what adding ISR does to the sdf. So all theorems in this section relate to T matrices and speak to existence and/or uniqueness of an equilibrium or an sdf, or spell out when T might have AOA and/or CM. In the case of the RA, we will also examine equilibrium stability.

To allow us to focus on the non-linearities introduced by systemic risk, this section adds the assumption of constant returns to scale, plus some minor assumptions (Recall that constant returns to scale was one of several optional assumptions required for sections 1.2 and 1.3). Also, the risk-free rate will continue to be endogenous, but some specification of the bond market is required: For this section, I assume that the risk-free asset, being a purely financial asset, is in zero net supply. There is no risk-free technology. Another change for this section is that returns are total returns, not measured in excess of risk-free. These assumptions mean that the sdf is not truly exogenous. While it does vary based on the utility function and initial wealth, it only varies because of the endogenous risk-free rate.

S is any NxN matrix of unimpacted asset returns which would yield a unique equilibrium without ISR added to it, for all agents with increasing concave utility. Using S as a baseline, I will add ISR and examine the existence and uniqueness of the equilibria of (5) - (9).

The fundamental theorems of asset pricing are useful Harrison and Kreps (1979); Harrison and Pliska (1981); Delbaen and Schachermayer (1994). We will find economically interesting violations of these theorems for the representative agent when ISR is present. The versions of the fundamental theorems that I use are:

In a well-defined asset market Harrison and Pliska (1981); Delbaen and Schachermayer (1994):

First theorem:

$$Absence\ of\ Arbitrage\ (AOA) \Leftrightarrow \exists\ SDF \equiv \mathbf{m}, s.t. E(\mathbf{m}R_i) = 1 \forall i$$

Second theorem (the theorem assumes AOA true throughout):

$$AOA: \ \ Complete\ Markets\ (CM) \Leftrightarrow \exists!\ \mathbf{m}, s.t. E(\mathbf{m}R_i) = 1 \forall i$$

Definitions of AOA and CM for asset market K:
$K, AOA \Leftrightarrow \nexists \boldsymbol{\beta} \in \mathbb{R}^{N-1}, \boldsymbol{\beta}'\mathbf{1} = 0,\ s.t.\ \hat{K}\boldsymbol{\beta} \succ^F \mathbf{0}$
$K, CM \Leftrightarrow \nexists \boldsymbol{\beta} \in \mathbb{R}^N, \boldsymbol{\beta} \neq \mathbf{0},\ s.t.\ K\boldsymbol{\beta} = \mathbf{0}$

Arbitrage (Absence of Arbitrage) is the existence (nonexistence) of some zero-cost portfolio of assets that first-order stochastically dominates a payoff of zero in every state of nature.

The symbol, $\succ^F$, means first-order stochastically dominates (FOSD). Asset $\mathbf{a}$ FOSD asset $\mathbf{b}$ ($\mathbf{a} \succ^F \mathbf{b}$) whenever $\mathbf{a}$ pays off at least as much as $\mathbf{b}$ in every state and pays an amount strictly greater in at least one state: $\mathbf{a} \succ^F \mathbf{b} \Leftrightarrow \mathbf{a} \geqq \mathbf{b}, \exists i, a^i > b^i$. In the definition for AOA, we may think of $K\boldsymbol{\beta}$ as a vector of portfolio payoffs from an arbitrage portfolio (a costless portfolio), with$\boldsymbol{\beta}$ representing the amount of wealth allocated to each asset. Portfolio allocations sum to zero for arbitrage portfolios. The risk-free asset is left out of the matrix in the definition for AOA above, so that $\boldsymbol{\beta}$ lives in $\mathbb{R}^{N-1}$. This is nonstandard. As a general rule, risk free assets can be included in arbitrage and the definition would have $\boldsymbol{\beta} \in \mathbb{R}^N$. For reasons discussed below, the risk-free asset is left out for our application, without loss of generality (wlog).

Complete markets are asset markets that can provide the investor any desired payoff profile state-by-state, with the purchase of some portfolio. The assets span the states. For us, this means the matrix under consideration is non-singular. In the definition for CM, we may think of $K\boldsymbol{\beta}$ as a vector of portfolio payoffs, with$\boldsymbol{\beta}$ representing the amount of wealth allocated to each asset. The sum of the components represent total allocated wealth and can be any real number.

The risk-free asset is not affected directly by the (stochastic) ISR. The risk-free rate is endogenous and cannot be specified in the asset market matrix S. Knowing R_f requires knowing the equilibrium solution, which also depends on the wealth and utility function. The theorems in this section apply to any permissible wealth and utility. When examining whether the ISR-impacted version of the available assets exhibits complete markets, it is sufficient to leave the risk-free column as a column of one's. This is because the question of complete markets is a question of spanning and does not depend on magnitude.

Regarding AOA, we know that the risk-free asset cannot be involved in arbitrage because its equilibrium holding is zero. Arbitraged assets do not have an equilibrium holding as agents essentially want an infinite (or negatively infinite) amount. If there is an arbitrage, which eliminates the possibility of equilibrium, it is not because of the risk-free asset. When considering arbitrage, the risk-free asset is left out of the S and T matrices and $\hat{S}$ and $\hat{T}$ are used instead. Whatever equilibrium R_f is settled upon, $T, AOA \Leftrightarrow \hat{T}, AOA$. This holds with the bond market that was specified, or with any specification for the bond market that has finite equilibrium holdings. The specified bond market was R_f endogenous and bonds in zero net supply.

2.2.2 Social Planner

During her optimization, the SP aggregates all the systemic risk externalities to $\alpha_A f(R_A)$ and assigns them to A. She effectively faces the following asset market with the effective asset returns shown. Assuming the first column is asset A and the last is the risk-free:

$$\left\{ \begin{array}{ccccc} s_1^1 - f(s_1^1) & s_2^1 & \dots & s_{N-1}^1 & 1 \\ s_1^2 - f(s_1^2) & s_2^2 & \dots & s_{N-1}^2 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ s_1^N - f(s_1^N) & s_2^N & \dots & s_{N-1}^N & 1 \end{array} \right\} \equiv T$$

Unlike the RA below, these effective assets are unchanging in $\boldsymbol{\alpha}$, so that the fundamental theorems of asset pricing apply. The entries in this matrix and the S matrix are total returns, not returns in excess of R_f as in all prior sections. All proofs for this section are at the end of the section.

Theorem 3. ***In the absence of arbitrage, markets are complete*** *(*$\forall S \in Q$, $T, AOA \Rightarrow T, CM$ *).*

The proof is robust to the equilibrium value of R_f, will be true regardless of equilibrium reached. The proof will rely on S, AOA (in statement 2), but not S, CM. We will now worry only about arbitrage.

Lemma 4. ***There are no cases of multiple SDF's.*** *(*$\nexists S.s.t. T, (\exists \mathbf{m}, \nrightarrow \exists! \mathbf{m})$*).*

The SP faces a unique sdf, or none. The proof is short as this plainly follows from the first fundamental theorem.

Lemma 5. ***A unique SDF is necessary and sufficient for a unique equilibrium*** *(*T, $\exists! \mathbf{m} \Leftrightarrow T$, $\exists! \boldsymbol{\alpha}^*$*).*

With unique pricing, the SP always finds a unique solution. This bodes well for stability of a Pareto optimal.

Corollary 5. ***There are no asset markets with multiple equilibria.*** *(*$\nexists S.s.t. T$, ($\exists \boldsymbol{\alpha}^*$, $\nrightarrow \exists! \boldsymbol{\alpha}^*$)

Obvious corollary from the above two lemmas. The agent finds one equilibrium, or fails to find any - the case we will worry about, if such a case exists.

To collect and summarize everything we have established in the above four Lemma / theorems, the only categories of T for $S \in Q$ that remain are:

1. $T:$ AOA, CM, $\exists! \mathbf{m}$, $\exists! \boldsymbol{\alpha}^*$
2. $T: \nrightarrow AOA$, $\nexists \mathbf{m}$, $\nexists \boldsymbol{\alpha}^*$

We do not have to consider incomplete markets, because they can only happen when their is arbitrage. Cases with multiple m or multiple equilibrium are impossible. We have not actually established that case 2, arbitrage, is possible - just that it is the only option to case 1 if it is possible. We prove in the next Lemma that it is possible, and in the Corollaries that follow Lemma 6, characterize its likelihood. Is it a common case, a knife-edge case, etc.?

Lemma 6. ***There are asset markets with no equilibrium.***

We prove that there is a case with arbitrage, the absence of any pricing kernel, and the absence of an equilibrium.

Corollary 6. ***The only asset markets with no equilibrium are those where the*** SP ***tries to sell*** A ***infinitely short.***

First, A's involvement is required for arbitrage. Removing asset A (column one) causes AOA for all cases. Secondly, there is no arbitrage that includes buying A, so the only one involves selling A. Another way to write this is: there is an arbitrage buying a convex combination of the other columns and selling column one, asset A.

Corollary 7. ***The only asset markets with no equilibrium are those where*** S ***itself is 'close' to arbitrage.***

Referring to the definition of arbitrage $\boldsymbol{\beta'}\mathbf{1} = 0,\ s.t.\ \hat{K}\boldsymbol{\beta} \succ^F \mathbf{0}$, I show that S is itself within $\beta^1\mathbf{f(s_1)}$ of arbitrage for $T, \nrightarrow AOA$. This only occurs for $\beta^1 < 0$.

2.2.3 Aggregate Agent

In this case, the returns of the assets are a function of the amount allocated to firm A, and the representative agent does not consider this.

$$\left\{\begin{array}{cccccc} s_1^1 & s_2^1 - \gamma_A f(s_2^1) & \dots & s_{N-1}^1 - \gamma_A f(s_{N-1}^1) & 1 \\ s_1^2 & s_2^2 - \gamma_A f(s_2^2) & \dots & s_{N-1}^2 - \gamma_A f(s_{N-1}^2) & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ s_1^N & s_2^N - \gamma_A f(s_2^N) & \dots & s_{N-1}^N - \gamma_A f(s_{N-1}^N) & 1 \end{array}\right\} \equiv T$$

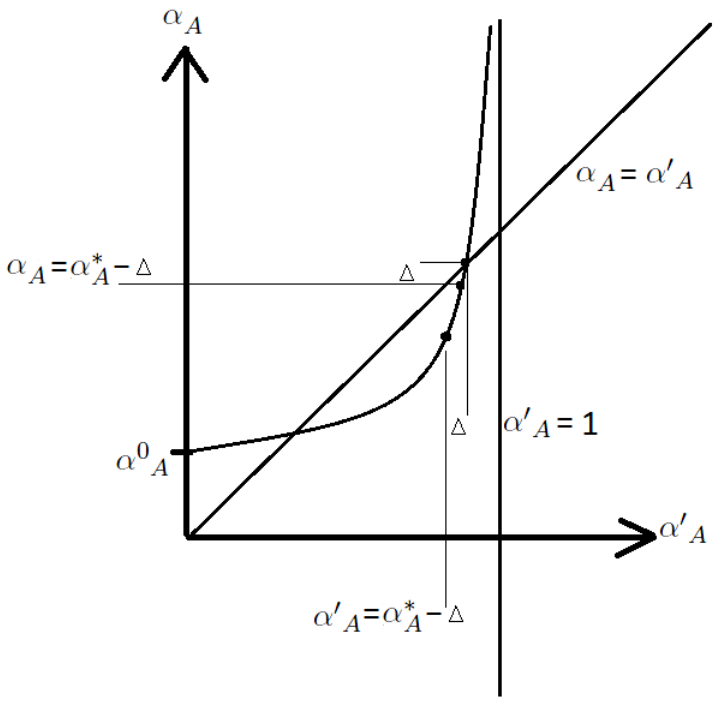
Where $\gamma_A = \frac{\alpha'_A}{1-\alpha'_A}$ and ${\alpha'}_A$ are treated by the agent as exogenous. Also, $s_1^i \equiv s_A^i$ is the systemic institution. Define ${\alpha^0}_A$ as the allocation to firm A in the absence of systemic risk. For any α_A to be an equilibrium α_A^*, we must have a fixed point such that the RA is optimizing at the returns she faces. In other words, we must have a fixed point such that $\alpha_A = \alpha'_A$. This can be seen in the graphic as the 45^o line.

Except for a knife-edge case of tangency, there are zero, or at least two, equilibria for the RA. When it's two, the "good" one is close to the equilibrium without ISR and the other is heavily invested in A, which is rational because heavy investment in A causes the other assets to have very low returns and be poor outside options. The latter equilibrium is also unstable, in that a deviation in allocation from the optimum results in the optimum

moving in a way that reinforces the deviation. The good equilibrium is stable. It is also a pareto improvement.

In the graphic, when the exogenous α'_A is zero, the RA would pick her $\alpha_A = \alpha^0{}_A$. This would not be possible because in the end, of course, $\alpha_A = \alpha'_A$ as the 45^o line shows. The instability of the upper equilibrium is demonstrated by a case where the RA deviates downward in her allocation to A by Δ, resulting in α'_A also decreasing by Δ, but this gives a new optimum well below the previous optimum of α^*_A, and even below $\alpha^*_A - \Delta$. The RA heads for this new optimum. A similar instability occurs if she deviates upward. The lower equilibrium does not have this problem. Deviations cause the RA to head back toward the optimum.

These portfolio adjustments are not technically part of the one-period model in this paper, but it is easy to imagine an multiperiod optimization with return realizations followed by rebalancing giving us a stability defined as I have. An important assumption would be the lack of fees for reallocation of capital, such as in Duffie and Strulovici (2012).



2.3 Part 2 Conclusion

The existence or uniqueness of a stochastic discount factor (sdf) generally follows the fundamental theorems of asset pricing Harrison and Kreps (1979); Harrison and Pliska (1981); Delbaen and Schachermayer (1994) regarding prices faced by the social planner. In line with our intuition, for the model presented, any asset market that would have a

unique sdf (and equilibrium) without ISR, would have them with ISR too - except in cases where the impacting firm creates its own arbitrage.

With RA pricing, the fundamental theorems of asset pricing cannot be used to determine the existence or uniqueness of a stochastic discount factor or equilibrium. It is violation of the fundamental theorems that makes systemic risk what it is. In the case of the representative agent, the optimum changes with her allocation, such that there is a notion of stability for asset market equilibria. When an equilibrium exists, at least two do. One of them is heavily invested in the systemically risky firm or industry, which is rational because heavy investment in it causes the other assets to face a large potential externality and be a poor outside option. The latter equilibrium is also unstable, in that a deviation in allocation from the optimum causes the optimum to move in a way that reinforces the deviation. The good equilibrium is stable.

2.4 Part 2 Proofs

Theorem 3: In the absence of arbitrage, markets are complete ($\forall S \in Q, T, AOA \Rightarrow T, CM$).

The proof is robust to the equilibrium value of R_f, will be true regardless of equilibrium reached. The proof will rely on S, AOA (in statement 2), but not S, CM. Prove the contrapositive ($T, \nrightarrow CM \Rightarrow T, \nrightarrow AOA$). $T, \nrightarrow CM$ implies that the first column of T equals a linear combination of the columns of the sub-matrix, $T^{\setminus 1}$, created by removing the first column from T:

$$t^1 = s^1 - f(s^1) = T^{\setminus 1}\boldsymbol{\beta} \tag{20}$$

I will prove two statements about the sum of the coefficients of this linear combination (about $\boldsymbol{\beta}^{\mathbf{T}}\mathbf{1}$ for the $\boldsymbol{\beta}$ in (20)) that will together prove the contrapositive result ($T, \nrightarrow AOA$).

1. $\boldsymbol{\beta}^{\mathbf{T}}\mathbf{1} \neq 1 \Rightarrow T, \nrightarrow AOA$

If the sum is not one, there is arbitrage for T. First consider $\boldsymbol{\beta}^{\mathbf{T}}\mathbf{1} < 1$. We can define a vector $\boldsymbol{\delta}$ that is a scaled up version of $\boldsymbol{\beta}$ and has components that sum to one ($\boldsymbol{\delta}^{\mathbf{T}}\mathbf{1} = 1$):

$$\boldsymbol{\delta}^{\boldsymbol{T}} \equiv \frac{\boldsymbol{\beta}^{\mathbf{T}}}{\boldsymbol{\beta}^{\mathbf{T}}\mathbf{1}}, \& \, t^1 = T^{\setminus 1}\boldsymbol{\beta} \Rightarrow T^{\setminus 1}\boldsymbol{\delta} >> t^1 \Rightarrow T^{\setminus 1}\boldsymbol{\delta} \succ^F t^1 \Rightarrow T^{\setminus 1}\boldsymbol{\delta} - t^1 \succ^F \mathbf{0} \Rightarrow T, \nrightarrow AOA$$

Where $\boldsymbol{\delta}^{\mathbf{T}}\mathbf{1} = 1$ means that the portfolio $T^{\setminus 1}\boldsymbol{\delta} - t^1$ is an arbitrage portfolio with weights that sum to zero. If $\boldsymbol{\beta}^{\mathbf{T}}\mathbf{1} > 1$, then $\boldsymbol{\delta}$ defined the same way is a scaled down version of $\boldsymbol{\beta}$, and t^1 is the dominant asset instead: $t^1 \succ^F T^{\setminus 1}\boldsymbol{\delta}$. In general, another way to think of arbitrage is one portfolio FOSD another in returns.

2. $\boldsymbol{\beta}^{\mathbf{T}}\mathbf{1} \neq 1$.

This is because $\boldsymbol{\beta}^{\mathbf{T}}\mathbf{1} = 1$ would imply $S, \nrightarrow AOA$, which is a contradiction by the definition of S. Assume $\boldsymbol{\beta}^{\mathbf{T}}\mathbf{1} = 1$ throughout:

$$s^1 - f(s^1) = T^{\setminus 1}\boldsymbol{\beta} \Rightarrow s^1 \succ^F T^{\setminus 1}\boldsymbol{\beta} \Longrightarrow s^1 \succ^F S^{\setminus 1}\boldsymbol{\beta}, \& \, \boldsymbol{\beta}^{\mathbf{T}}\mathbf{1} = 1 \Longrightarrow S, \nrightarrow AOA$$

Where we used that $f(s^1) \succ^F \mathbf{0}$ and $T^{\setminus 1} = S^{\setminus 1}$.

Lemma 4: There are no cases of multiple SDF's. $(\nexists S.s.t.T, (\exists \mathbf{m}, \multimap \exists! \mathbf{m}))$.

Definition SDF, $\mathbf{m} : \mathbf{m} \in \{\mathbf{x} : E(\mathbf{xt_i}) = 1 \forall i\}$.

if AOA : $AOA \Rightarrow CM$ by Theorem 3. $(AOA,CM) \Leftrightarrow \exists! \mathbf{m}$ by second fundamental theorem.

if $\multimap$ AOA: Arbitrage eliminates all $\mathbf{m}$ and equilibria.

Lemma 5: A unique SDF is necessary and sufficient for a unique equilibrium $(\exists! \mathbf{m} \Leftrightarrow \exists! \boldsymbol{\alpha}^*)$.

Case1: $T, \multimap AOA$. Arbitrage eliminates all $\mathbf{m}$ and equilibria, consistent with the lemma.

Case2 T, AOA. Assume AOA throughout this proof. Theorem 3 is $T, AOA \Rightarrow T, CM$. So we also have CM throughout. Hence we may also assume $\exists! \mathbf{m}$ throughout by the second fundamental theorem. Do we have a unique equilibrium?

The optimization problem can be written $max_{\boldsymbol{\alpha}} E(u(T\boldsymbol{\alpha}))$ $s.t.$ $\sum \alpha_i = 1$, where $\boldsymbol{\alpha} \in \mathbb{R}^N$. T, CM implies that there is a unique set of prices for the N Arrow-Debreu securities which represent an equivalent investment set as T Arrow and Debreu (1954). Rewrite the optimization as $max_{\boldsymbol{\alpha}} E(u(\sum \alpha_i A_i))$ $s.t.$ $\sum \alpha_i = 1$, where A_i is the return of the Arrow-Debreu security paying one dollar in state i and zero in all other states, and α_i is the allocation to that security. If any solutions that exist are interior solutions only, and the objective function is increasing and strictly concave, then there exists a unique solution to this maximization problem per De la Fuente (2000). Permitting shorting of the A_i's implies interior solutions only. For the objective function to be strictly concave, $\forall \boldsymbol{\alpha^1}, \boldsymbol{\alpha^2} \in \mathbb{R}^N$, $\delta^1, \delta^2 \in \mathbb{R} : \delta^1 + \delta^2 = 1$ we must have:

$$Eu((\delta^1 \sum \alpha_i^1 A_i + \delta^2 \sum \alpha_i^2 A_i)) > \delta^1 E(u(\sum \alpha_i^1 A_i) + \delta^2 E(u(\sum \alpha_i^2 A_i)$$

$$\Leftrightarrow \sum \pi_i u((\delta^1 \alpha_i^1 A_i + \delta^2 \alpha_i^2 A_i)) > \sum \pi_i (\delta^1 u(\alpha_i^1 A_i) + \delta^2 u(\alpha_i^2 A_i))$$

Which is true by the strict concavity of $u(\cdot)$: $u((\delta^1 \alpha_i^1 A_i + \delta^2 \alpha_i^2 A_i)) > \delta^1 u(\alpha_i^1 A_i) + \delta^2 u(\alpha_i^2 A_i)$, $\forall i = 1, N$.

Corollary 5: There are no asset markets with multiple equilibria.

if $\multimap$ AOA: Arbitrage eliminates all $\mathbf{m}$ and equilibria.

if AOA : $AOA \Rightarrow CM$ by Theorem 3. $(AOA,CM) \Leftrightarrow \exists! \mathbf{m}$ by second fundamental theorem. Combining this with Lemma 5 implies that $AOA \Leftrightarrow \exists! equilibirum$.

Lemma 6: There are asset markets with no equilibrium.

As the first fundamental theorem of asset pricing applies, it will be sufficient (and necessary) to have $\exists T, \nrightarrow AOA$ to prove no SDF, which by Lemma 5 would prove this lemma. Recalling that the last column will not be involved in arbitrage, we will discard the last column of S and T, and work with the resulting submatrices $\hat{S}$ and $\hat{T}$: $T, AOA \Leftrightarrow \hat{T}, AOA,\ S, AOA \Leftrightarrow \hat{S}, AOA$. Let all $\boldsymbol{\beta}$ in this proof have $\boldsymbol{\beta} \in \mathbb{R}^{N-1}$, $\boldsymbol{\beta}^{\mathbf{T}}\mathbf{1} = 0$ and $\beta^1 < 0$:

No matter how small the epsilon ($\boldsymbol{\epsilon} \in \mathbb{R}^{N-1},\ \boldsymbol{\epsilon} \succ^F \mathbf{0}$):

$$\forall \boldsymbol{\epsilon},\ \exists S \in Q,\ s.t.\ \hat{S}\boldsymbol{\beta} \succ^F -\boldsymbol{\epsilon} \tag{21}$$

for some $\boldsymbol{\beta}$.

Because $\mathbf{0} \succ^F -\boldsymbol{\epsilon}$, this does not violate S, AOA. The S in (21) exists because Q contains all asset markets S, which by definition have AOA, even those close to arbitrage.

Now set $\boldsymbol{\epsilon} \doteq \hat{T}\boldsymbol{\beta} - \hat{S}\boldsymbol{\beta} = -\beta^1\mathbf{f}(\mathbf{s_1}) \succ^F \mathbf{0}$. With this epsilon, choose an S and $\boldsymbol{\beta}$ that fulfill (21):

$$\hat{S}\boldsymbol{\beta} \succ^F -\boldsymbol{\epsilon}\,(= \hat{S}\boldsymbol{\beta} - \hat{T}\boldsymbol{\beta}) \ \Rightarrow\ \hat{T}\boldsymbol{\beta} \succ^F \mathbf{0} \Leftrightarrow T, \nrightarrow AOA$$

Corollary 6: The only asset markets with no equilibrium are those where the SP tries to sell A infinitely short.

First, A's involvement is required for arbitrage. This is because $S, AOA \Rightarrow \hat{S}^{\setminus 1}, AOA$, *and* $\hat{S}^{\setminus 1} = \hat{T}^{\setminus 1} \Rightarrow \hat{T}^{\setminus 1}, AOA$. So there is no arbitrage if column 1 is excluded, i.e. if $\beta^1 = 0$. The two steps below are: 1. Prove that there is an arbitrage where the investor shorts A. 2. Prove there is no other kind of arbitrage.

1. Let all $\boldsymbol{\beta}$ in both parts of this proof have $\boldsymbol{\beta} \in \mathbb{R}^{N-1}$, $\boldsymbol{\beta}^{\mathbf{T}}\mathbf{1} = 0$. For this part 1 of the proof, $\beta^1 < 0$. Define $\boldsymbol{\delta} \equiv \frac{\boldsymbol{\beta}^{\setminus 1}}{-\beta^1} \Rightarrow \boldsymbol{\delta}^{\mathbf{T}}\mathbf{1} = 1$:

$$T, \nrightarrow AOA \ \Leftrightarrow\ \hat{T}\boldsymbol{\beta} \succ^F \mathbf{0} \ \Leftrightarrow\ \hat{T}^{\setminus 1}\boldsymbol{\beta}^{\setminus 1} \succ^F -\beta^1\mathbf{t_1}$$

$$\Rightarrow\ \hat{T}^{\setminus 1}\boldsymbol{\delta} \succ^F \mathbf{t_1},\ \&\, \boldsymbol{\delta}^{\mathbf{T}}\mathbf{1} = 1$$

This represents an arbitrage where asset $\mathbf{t_1}$ is shorted against some other portfolio.

2. Step 1 shows the case where $\beta^1 < 0$, leaving $\beta^1 > 0$. Because $\mathbf{s_1} = \mathbf{t_1} + \mathbf{f}(\mathbf{s_1})$ and $\hat{T}^{\setminus 1}\boldsymbol{\beta}^{\setminus 1} = \hat{S}^{\setminus 1}\boldsymbol{\beta}^{\setminus 1}$, $\hat{S}\boldsymbol{\beta} = \hat{T}^{\setminus 1}\boldsymbol{\beta} + \beta^1\mathbf{t_1} + \beta^1\mathbf{f}(\mathbf{s_1}) = \hat{T}\boldsymbol{\beta} + \beta^1\mathbf{f}(\mathbf{s_1})$:

$$\hat{S}, AOA \Rightarrow \nrightarrow (\,\hat{T}\boldsymbol{\beta} + \beta^1\mathbf{f}(\mathbf{s_1}) \succ^F \mathbf{0}\,) \Rightarrow \nrightarrow \hat{T}\boldsymbol{\beta} \succ^F \mathbf{0} \Leftrightarrow T, AOA$$

Where we used that $\beta^1 > 0$ and $\mathbf{f}(\mathbf{s_1}) \succ^F \mathbf{0}$.

Corollary 7: The only asset markets with no equilibrium are those where S itself is 'close' to arbitrage.

Being the systemic externality, $f(s_1^i)$ is near zero for most reasonable arguments and positive for low ones. The magnitude of β^1 is less than one (by the assumption on α_A),

possibly much less if there are many assets. From above, $\hat{T}\boldsymbol{\beta} = \hat{S}\boldsymbol{\beta} - \beta^1\mathbf{f}(\mathbf{s_1})$.

$$S, AOA \Leftrightarrow \rightharpoonup \hat{S}\boldsymbol{\beta} \succ^F \mathbf{0}\,, \quad T, AOA \Rightarrow \rightharpoonup \hat{S}\boldsymbol{\beta} \succ^F \beta^1\mathbf{f}(\mathbf{s_1})$$

$$T, \rightharpoonup AOA \Leftrightarrow \mathbf{0} \succ^F \hat{S}\boldsymbol{\beta} \succ^F \beta^1\mathbf{f}(\mathbf{s_1})$$

$\hat{S}$ is within $\beta^1\mathbf{f}(\mathbf{s_1})$ of arbitrage.

3 Arbitrage Pricing Theory and Systemic Risk

3.1 Introduction

This Part presents a factor model based on the arbitrage pricing theory (A.P.T.) model and the model in Part 1. Everything being written in terms of returns brings ISR solidly into the domain of *systematic* risk (SR). I use the model to distinguish between systematic and systemic risks. The model shows why systematic risk is so often mistaken as systemic risk, why systematic risk in the financial industry is important, and why it should be considered along with systemic risk in regulatory efforts. The model is then used to delineate and outline the various types of risk. This vocabulary can facilitate communication and research in systemic risk.

This kind of factor model varies from prior specifications of returns (even prior specifications of returns that included SR and ISR) in two important ways. First, the factor model is an asset-pricing model, rather than a return-generating specification. Secondly, the factor model contains ISR as systemic risk, rather than simply adding non-Gaussian idiosyncratic or systematic terms as proxies for systemic risk. Any distorted pdf's that appear will be the result of ISR not ISR itself. Furthermore, correlated tails appear in the returns of the *impacted* institutions, raising questions about which firms to regulate - the fact that the correlated tails might be found in the firm being affected by the systemic externality, rather than in the firm causing it, is important.

Recall that unimpacted returns (R') are the returns that would have happened in the absence of ISR: $R' = R - f(\cdot)$. The pricing for the impacted firms is derived in terms of the unimpacted returns and in terms of A.P.T. pricing. They are shown to be equivalent. The ISR decomposes into additional systematic risk, additional mean-zero idiosyncratic risk, and a constant negative externality - all on the impacted firm.

As one of the earlier attempts to make systemic risk and asset prices endogenous, the model is simple by asset pricing standards. It is also simpler than many attempts to characterize returns, such as the copulas and joint return topologies in quantitative finance, as well as highly technical efforts to use assumed fat-tailed returns to derive allocation and pricing. The model is also simpler than many attempts to characterize returns, such as the copulas and joint return topologies in quantitative finance, as well as highly technical efforts to use assumed fat-tailed returns to derive allocation and pricing.

The difference is that the model presented here has the potential to derive the implications of systemic externalities for the pdf's of returns rather than measure or characterize them.

3.2 Assumptions and Model

The assumptions and model follow. Throughout **A1 - A3**, and the results derived, the model is a full factor model with any shape for the idiosyncratic risk (Gaussian, fat tails, etc.), any shape for the systematic risk (Gaussian, fat tails, correlated tails, etc.), and any function for the systemic risk. After that, additional optional assumptions are added. If we think of this as a one-period model, $f(R_i)$ can be viewed as negatively affecting the ending cash value of the systemically impacted firm with no further interpretation needed[11]. In a multiperiod model, the ISR term $f(R_i)$ can be viewed as a causal reduction (over one time-step) to the expected sum of sdf-discounted life-time cash-flows of the firm - in other words, $f(R_i)$ is one part of the Campbell-Shiller innovation, see Campbell and Shiller (1988), over the time-step[12]. This model easily generalizes to many systemically important firms, many systemically impacted firms, and many firms that are both. Later, I add distributional assumptions which will allow future efforts to determine the distributional effects of ISR. In that model, generalizing to many firms is even more straightforward, as is shown.

The factor model:

A1. In the absence of systemic risk, returns follow a factor model.

$$R_i' = k_i + b_i' I + \varepsilon_i = m_i + b_i' g + \varepsilon_i \; \forall i$$

As is common, the random vector of (linearly independent) factors can be redefined as deviation from expectation, $g = I - E(I)$, $E(g) = E(\varepsilon_i) = 0$. The vector of loadings is denoted by b_i, and k_i and m_i are constants. The assumption applies to systemically important and systemically impacted firms. This means that in the absence of systemic risk, the stochastic component of returns comes from loading on systematic priced factors plus a mean-zero, independent, firm-specific shock ε_i. Pricing, however, will be determined by the A.P.T. model.

A2. For all assets, the A.P.T. assumptions hold, per Ross (1976). For some asset i

$$R_i = E(R_i) + b_i' g + \varepsilon_i \,,\; E(R_i) = b_i' \lambda, \;\; \Rightarrow R_i = b_i' \lambda + b_i' g + \varepsilon_i$$

where λ are risk premia. Where the assumption of a factor model implies the first equation and A.P.T. pricing implies the second. All returns are in excess of R_f. λ, c, b, $\hat{b}$

[11] Although information is not modeled directly here, we should view it as a full information model, where all innovations are realizations of random variables rather than discovery of information.

[12] The reduction to the expected sum of discounted cash-flows can come from two things: the first is a change to expectations about cash flows and how they vary by state (resulting in changes to the expected value of cash and/or to the covariance of cash with the sdf), the second is a change in how the sdf varies by state. The first represents the cash-flow channel, and the second represents the discount factor channel. Because of the question this paper is asking, these channels are not modeled explicitly; the functional form would be determined by the particular ISR model being examined, but several results can be obtained even for a general form. Most ISR theories focus on the cash-flow channel. See Bai et al. (2012), and others, for papers with counterparty ISR that acts through both cash-flow effects and by moving the sdf - such models could still be accommodated in this framework to derive their implications for distributions and OSR.

are vectors; $f(R_i)$, ε, R, and x are random variables; g is a random vector; i, j specify assets; k specifies a factor from here going forward. On R, the prime symbol (as in R') indicates unimpacted return as before. On b and c it is the vector's transpose.

A3. As with the main model, ISR is a function of the returns of the systemically important firm and is reduced from the returns of systemically impacted firms: $R_j = R'_j - f(R_i)$, for all impacted firms j, where R'_j is the unimpacted return of firm j.

Recall that R'_j is the unimpacted return of asset j, one of the systemically impacted assets. $R_j = R'_j - f(R_i)$. This is an issue. If the returns R'_j follow A.P.T pricing, it is unlikely that R_j will too. This is because subtracting the random variable $f(R_i)$ changes loading and expected value, and probably will not do so in a way that maintains correct pricing. The change in loading on factors changes the required A.P.T. expected return.

3.2.1 Basic Pricing of the Systemically Impacted Firms

To solve this, return to the cash-flows: x_j is firm j's actual ending cash value, x'_j is the unimpacted ending cash value (what the value would be if the firm were not impacted by any ISR externality - for most realizations they are the same). The price, as well as the allocation of the RA or SP, is α_j; and it must follow A.P.T. pricing. The price will be determined using x_j, actual ending cash value, because the assumption is that all assets follow A.P.T. pricing, and the unimpacted asset with ending cash value of x'_j is a counterfactual asset.

$$x'_j = E(x'_j) + c'_j g + \epsilon_j \Rightarrow R'_j = E(R'_j) + b'_j g + \varepsilon_j \quad \not\Rightarrow \quad R'_j = b'_j \lambda + b'_j g + \varepsilon_j$$

where b_j has been defined here as the loading vector for the unimpacted returns. The unimpacted cash generated by firm j is a constant plus factor loadings and a firm-specific shock. The assumption in this paper is that the cash-value of ISR is a percentage of the firm: $-\alpha_j f(R_i)$.

$$x_j = x'_j - \alpha_j f(R_i) \Rightarrow R_j = E(R'_j) + b'_j g + \varepsilon_j - f(R_i)$$

Furthermore, we know that α_j, the initial price of asset j, will be such that R_j follows A.P.T. pricing, a fact we have yet to use.

3.2.2 Arbitrage Pricing of the Systemically Impacted Firms

Let the number of states be W and the number of factors be N. Let Y be the subspace of $\mathbb{R}^W$ (or of the corresponding infinite-dimensional space if there are an infinite number of states) spanned by g. As idiosyncratic shocks exist, we know that g does not span all of W. Decompose $f(R_i)$ into the three parts. First, there is the portion that can be projected onto the factors, onto Y; call this portion $f(R_i)_Y$. $f(R_i)$'s projection has created a projection matrix, which is a random vector. Because this projection matrix lives in Y, it can be written as a vector of constants times the random variables that comprise g. Let this N-dimensional constant be e_j. R'_j has factor loadings b_j, and now R_j has factor loadings

$b_j + e_j$, which we will call $\hat{b}_j$. The remaining component of $f(R_j)$ is $[\, f(R_i) - f(R_i)_Y \,]$. This remaining component can be decomposed into two parts. There is a constant term, $E[f(R_i) - f(R_i)_Y\,]$, and mean-zero noise, $\{\, [f(R_i) - f(R_i)_Y\,] - E[\, f(R_i) - f(R_i)_Y\,]\}$. All three of these components of $f(R_i)$ will appear in our equation. From A.P.T. pricing:

$$R_j = R'_j - f(R_i) = E(R'_j) + b'_j g - f(R_i) + \varepsilon_j = E(R'_j) + \hat{b}'_j g - [\, f(R_i) - f(R_i)_Y\,] + \varepsilon_j$$

$$= \{E(R'_j) - E[\, f(R_i) - f(R_i)_Y\,]\} + \hat{b}'_j g - \{\, [f(R_i) - f(R_i)_Y\,] - E[\, f(R_i) - f(R_i)_Y\,]\,\} + \varepsilon_j$$

$$= E(R_i) + \hat{b}'_j g + \eta_j \;=\; \hat{b}'_j \lambda + \hat{b}'_j g + \eta_j$$

where $\eta_j = [\, f(R_i) - f(R_i)_Y] \,-\, E[\, f(R_i) - f(R_i)_Y\,] + \varepsilon_j$, the noise part of $f(R_i)$ plus the noise part of R'_j.

Now define stochastic return shocks $\acute{R}_i \equiv R_i - E(R_i)$ and $\acute{R}_j \equiv\, R_j - E(R'_j)$.

$$\underbrace{\acute{R}_i = \underbrace{b'_i g}_{\text{systematic risk}} + \underbrace{\varepsilon_i}_{\text{idiosyncratic risk}}}_{\text{systemically important institution}}$$

$$\underbrace{\acute{R}_j = \underbrace{b'_j g}_{\text{systematic risk}} - \underbrace{f(R_i)}_{\text{systemic risk}} + \underbrace{\varepsilon_j}_{\text{firm-specific risk}}}_{\text{systemically impacted institution}} \tag{22}$$

and

$$R_i = b'_i\lambda + b'_i g + \varepsilon_i \qquad R_j = \; E(R'_j) \; + b'_j g - \; f(R_i) \; + \underbrace{\varepsilon_j}_{\text{firm-specific}} \; = \; \hat{b}'_j\lambda \; + \hat{b}'_j g + \underbrace{\eta_j}_{\text{idiosyncratic}}$$

The model has rigorously defined pricing, systematic risk, systemic risk, and firm-specific risk. It is used in Section 3.3 to clarify and delineate the various types of systemic risk.

Because R_i and g are correlated, but imperfectly so, the projection of the random variable $f(R_i)$ onto the priced factors is non-zero but of lower magnitude than $f(R_i)$. In other words, part of this risk is priced. Until now, firm-specific and idiosyncratic risks have been synonymous, but here we need to distinguish between them. I do not view this as an artifact of the model. If ISR exists, it imposes additional idiosyncratic risk. Part of firm j's total idiosyncratic risk, η_j, comes from the firm-specific risk of firm i being transmitted to firm j via ISR, $\{\, [f(R_i) - f(R_i)_Y\,] \,-\, E[f(R_i) - f(R_i)_Y\,]\,\}$, and part comes from the firm-specific risk of j itself, ε_j. All three of these are diversifiable, mean-zero, and unpriced. It is already clear that if unimpacted returns are normal, and $f(\cdot)$ is anything but affine, then impacted returns will not be normal. It should also be clear that the returns of j are correlated with the idiosyncratic risk of i, another non-standard asset-pricing result.

No results have been derived here. The model and pricing have just been laid out and offered for use in future work. If the optional assumptions are employed, this could

lead in future work to testable implications and signatures that we can look for in the returns of impacted firms, and in the joint returns of impacted and impacting firms. This depends upon whether the systemic risk information is in the returns - do the deviations from normality caused by ISR dominate the deviations from normality that may come from other sources?

3.3 Application: *The Various Types of Systemic Risk*

There is a large literature that assumes the salient aspect of systemic risk is the possibility of a financial crisis, which impacts the real economy. This sub-section stands within that large literature. The assumption is that the financial system may have very poor performance one period, this will result in the intermediation capacity of the financial system failing, and this will result in a large negative externality on the real economy. Finally, as a proxy for the event that a financial crisis occurs, this large literature (and this section) use some threshold value for the returns of the financial industry as a whole. Any return realization for the industry which is below this threshold represents crisis and the resulting market failure and welfare loss.

I am not so much strongly advocating for this view as working for its clean implementation. Only this kind of clarity can allow it to be tested.

In this view of things, regulators should care about anything that contributes to the chance of a crisis (the probability that the financial industry will have a return below the threshold), about the magnitude of the negative externality (per unit time) during crisis, and about the duration of crisis (see Ibragimov et al. (2011) for the importance of duration). As with the majority of this large literature, I do not model the magnitude or duration, but instead focus on the probability that the financial industry has an overall return below a constant threshold. I call this probability Overall Systemic Risk (OSR). One could think of OSR as the chance of a financial crisis occurring this year for example. Approximately following Acharya et al. (2010); Ibragimov et al. (2011) and many others, I define a *systemic financial crisis* as the event $\{\omega \in \Omega : R_I < \underline{R}\}$ and *overall systemic risk* (OSR) as $\mathbb{P}\{R_I < \underline{R}\}$, the chance of a financial crisis.

Wherever possible, I define terminology as it is already in use. In a few cases, terms are currently being used in the literature for more than one phenomenon per term, where it is not clear from context how it is being used. Such a scenario makes it necessary to create a new term. In other cases, risks that are well established and known in the asset pricing literature are being referred to as systemic risk in the systemic risk literature. In such cases, it is not necessary to create a term or even address it directly. Outlining each type of risk allows specific terms to be used. I am not inventing any new types of systemic risks in this paper. One point that the model makes clear is that the difference between systemic and systematic risks is the difference between causality and correlation.

With the assumptions that the magnitude and duration of the crisis' externality are constants, the planner should, *ceteris paribus*, seek to minimize OSR[13]. She would be

[13] Alternatively, several papers have a simple utility function that is dramatically impacted by a financial

concerned with anything that increases OSR - even if it does not represent a market failure or inefficiency in and of itself. The planner should also be concerned with anything that does represent a market failure and inefficiency in and of itself, even if it does not increase OSR. Finally, the planner would have two reasons to be concerned with any risk that increases OSR on the one hand, and is a market inefficiency in and of itself on the other.

The goal of this section is to identify and characterize, within this view of systemic risk, any risk or moral hazard that increases OSR, stands as a welfare-reducing market failure in and of itself, or both. The method of distinguishing these types of risks is to reference an Arbitrage Pricing Theory (A.P.T.) model of the assets in the financial industry shown in equation (22) above.

For example, the firms in the financial industry having a lot of systematic risk (high beta's, or in this case, b's) increases industry volatility and hence increases OSR. Even though the term "systemic risk" does not appear in the name of systematic risk, it could be considered to be a systemic risk only because it increases the chance of a financial crisis, i.e. it increases OSR.

Firms, even banks, increasing their loading on factors (factors such as the return of the market or the Fama-French factors, or any risk factor) is systematic risk, even though it increases their correlation and increases the chance of a financial crisis. The terms defined will also have use in other systemic risk literatures, such as the literature on contagion, but the paper should be viewed as living in the world described in the preceding paragraph. The terminology will not fit perfectly into all systemic risk literature, but will at least not conflict.

Types of risk that have systemic importance in this view of the world:

1. Inter-industry systemic risk. This is the externality itself that comes from the financial industry and affects the real industry. It is causal.

2. Overall systemic risk (OSR). The probability of a financial crisis, of inter-industry systemic risk being realized.

3. Systematic risk in the financial industry. Systematic risk is risk within a firm return's that is both non-causal and correlated the stochastic discount factor - and therefore undiversifiable. (Non-causal means there is nothing to model the fact that the amount of asset i's systematic risk is caused by the returns of asset j. Also, non-causal means that there is nothing in the model representing i's systematic risk as affecting other returns, the distribution of the factors, etc.) In a factor model, the amount of systematic risk in the financial industry is the overall magnitude of financial firms loading onto the systematic factors. The systematic factors do not need to be Gaussian. Systematic risk is only a problem in that it increases correlation of financial firms and industry variance and hence OSR.

crisis. Maximizing this utility function is very similar to, and usually identical to, minimizing OSR.

4. Interfirm systemic risk. This is a spillover, a risk imposed one one institution by another, a contagion, or a stochastic externality. It is causal. This is welfare-reducing in and of itself if it is not somehow priced. It also increases correlation of financial firms and industry variance and hence OSR.

5. Systemic allocation risk. The prospect of systemic risk causes sub-optimal allocation, even worse than the allocation we would have if the market ignored the systemic risk. In other words, not only does the market fail to take into account the systemic risk and adjust investment accordingly, the market actually reduces capital allocation efficiency when faced with systemic risk. As far as I know, this systemic risk was discovered (theoretically at least) in Part 1 of this dissertation.

6. Correlated factor risk, in the financial industry. For a given level of overall systematic risk, i.e. a given level of total loading on factors, there can be spreading of the loading of the various firms onto the various risk factors – or the firms can all herd and load-up on the same factor. Neither of these is a problem by itself, the market achieves efficient allocation in the face of this. But it increases OSR, and so it is a problem when combined with the possibility inter-industry systemic risk.

7. Idiosyncratic risk. Idiosyncratic risk is risk within a firm return's that is both non-causal and uncorrelated with any priced risk, and usually uncorrelated with any other risk at all. In a factor model, it is risk that is uncorrelated with factors, and hence other assets. Idiosyncratic risk does not need to be Gaussian. It generally would not increase OSR, unless the firm was very large and/or the firm had a very large potential interfirm systemic externality.

As the literatures which contain systematic risk co-mingles with the literature containing systemic risk, we are naturally in a phase where confusion about the two may arise. There is sometimes confusion about item 4 being causal. In contrast, there is never confusion about item 1 being causal. There can sometimes be confusion about whether item 3 is still item 3 when the pdf if not Gaussian.

Finally, there are two moral hazards:

1. Risk-taking moral hazard. If managers know that their firm might get bailed-out to stop inter-industry systemic risk, then they may take on more risk, such as systematic risk. They might also engage in factor-herding to make sure their firm fails when others do, and hence when bailouts are happening.

2. Allocation moral hazard. As mentioned, just the potential of inter-industry or interfirm systemic risk distorts capital allocation and increases allocation to the systemically risky firm or industry. If the prospect exists that the risky firm might be bailed-out ex post, then it should seem obvious that investors will likely want to invest even more in this firm, increasing the magnitude and possibility of interfirm systemic risk further. A less intuitive result is that allocation moral hazard can exist even if the investors or managers of the risky firm are not allowed to benefit from

the bailout, if the bailout is used only to reduce the systemic externality. As far as I know, this moral hazard was discovered (theoretically at least) in Part 1 of this dissertation.

With these definitions, we can characterize various types of risk, and what we should do about them:

Types of relevant risk:

ISR	Interfirm systemic risk
SR	Systematic risk
$IDIO$	Idiosyncratic risk
OSR	Overall systemic risk
CFR	Correlated factor risk
$IISR$	Inter-industry systemic risk

Actions:

A	Regulate in theory because it is an unpriced externality itself. Additional issue if we believe that the firm is in an industry whose poor performance could hurt the real economy.
B	Should not regulate, even in theory, *unless* we believe that the firm is in an industry whose poor performance could hurt the real economy.
C	Same as B but contributes via two pathways to increase OSR, making it potentially more impactful.
D	Do not regulate.

List format:

[Phenomenon] – [Type(s) of Risk], [Appropriate Action]

- Counterparty risk – ISR, A
- Margin spiral risk – ISR, A
- One firm holding credit default swaps written on another – ISR, A

- Financial firms levering up – SR, B
- Many financial firms exposing themselves to one type of risk, such as real-estate prices – CFR, B
- A bank with *intra*-industry systemic importance taking on too much risk – SR, C**
- Hedge funds manipulating stock prices (See Ben-David et al. (2013)) - ISR^*, A
- Highly correlated financial firms with heavy loading on factors – CFR, B
- Banks in general taking on too much risk, or the banking industry taking on too much risk – SR, B
- Financial firms all with heavy loading on particular factors that have fat tails, causing high tail correlation – CFR, B
- Externalities that are only operative when the impacting firm is in distress, causing high tail correlation – ISR, A
- Factors have fat tails, causing high tail correlation – SR, B
- A financial firm faces tipping points in its own profitability making firm-specific risk with fat tails – $IDIO$, D
- Contagion - ISR, $IISR^+$, A
- The two-way linkages between the banking sector and the macro-economy - $IISR$, A
- A bank taking too much risk - SR and/or $IDIO^{++}$, B
- Intermediation capacity of the system impaired – $IISR$, A

* To the extent that actual economic losses occur

** The bank is adding its own increase of SR, which increase volatility and OSR of the financial industry. It is also increasing its own propensity to cause ISR by becoming riskier, which is a welfare loss in its own right, and increases volatility and OSR of the financial industry. Furthermore, because the increase in the propensity to cause ISR is coming from an increase in SR, the increase in propensity is to cause ISR in bad states of nature, when the industry could be near financial crisis. This is because SR is, by definition, pro-cyclical.

$^+$ Or international systemic risk

$^{++}$ The statement is not fully defined, but the idiosyncratic part does not matter.

3.4 Potential Applications for Returns

To derive implications for joint return densities in future work, further assumptions are needed. I proceed with distributional assumptions.

A4-O. Optional Assumption for certain applications. There are many assets. In the absence of systemic risk, their returns are normal. This nearly follows observation for most firms. Furthermore, the results of this model may explain the deviation that *is* observed.

A5-O. Optional Assumption for certain applications. Firm-specific risks are independent and normal. I follow the empirical asset pricing work on factor models, which often assumes normal, independently distributed return errors. We have every reason to imagine that firm-specific risk, being the sum of many random variables, is Gaussian, see for example Casella and Berger (2001) and the Central Limit Theorem.

A6-O. Optional Assumption for certain applications. $f(\cdot)$, being the interfirm externality, is weakly positive, equal to zero for all positive arguments, and decreasing.

We still have:

$$R_i = b_i'\lambda + b_i'g + \varepsilon_i \qquad R_j = b_j'\lambda + b_j'g - f(R_i) + \underbrace{\varepsilon_j}_{\text{firm-specific}} = \hat{b}_j'\lambda + \hat{b}_j'g + \underbrace{\eta_j}_{\text{idiosyncratic}}$$

Again, ISR imposes additional idiosyncratic risk on firm j. Part is transmitted firm-specific risk of i , and part comes from the firm-specific risk of j itself. Although $E(\eta_j) = 0$, η_j is, importantly, not Gaussian and not independent. The implication of including ISR is that R_j is also not Gaussian. The use of only one firm causing the externality in this and prior models in this dissertation, and having only one impacted firm in the A.P.T. model, are often unnecessary. This is particularly true in this section. The upshot of the forms of returns is that R_i and R_j' are bivariate normal with non-zero correlation. This would still be true if i were affecting other firms or j was being affected by other firms. The firms' externalities could be weighted by parameters of each firm or of it's relationship to j, giving $-f(\sum_i w_i R_i)$, still an argument that is bivariate normal with R_j. Even the decomposition of the ISR into its three types of components would stand.

Sections 3.3, 3.4, and 3.5 utilize the arbitrage pricing theory model from Section 3.2. The model is simple by asset-pricing standards (this is as it should be; systemic risk work also strives to achieve other things), but the model a adds a few things not in most systemic risk work - hence, it has potential to add to the exciting work being done on systemic risk and returns. This brings up another potential application. Current systemic risk measures that are based on such intuition and use return panels for measuring systemic risk, such as marginal expected shortfall and CoVaR, can be assessed directly to shed light on what they measure. One could even solve for these measures in terms of the parameters of a pricing model now that a general causal externality is modeled in returns space. This can then be examined to see if we believe that the resulting expression captures ISR. This is done below in Section 3.5. The conclusion of Part 3 discusses other future efforts.

3.5 Derivation of Systemic Risk Measure from an Asset Pricing Model

CoVaR (Adrian and Brunnermeier (2011)) is a measure of a firm's contribution to overall systemic risk, based on the assumption that there should be evidence of ISR in the joint returns of impacted and impacting firms. The section above shows (from a generalized form of ISR written in returns space) that this may indeed be true. Does the particular form of the metric CoVaR capture these implications? CoVaR focuses on the financial institution and the financial industry. The factor model in Section 3.2 can be used to derive CoVaR and other systemic risk measures directly in terms of the parameters of a pricing model. I chose CoVaR because it has an interesting term which appears to be the result of deliberate attempts by Adrian and Brunermeier to capture spillovers or interfirm externalities in their measure. These attempts came only from economic intuition.

To briefly review CoVaR:

$\Delta CoVaR$ stands for Co-Value-at-Risk, and can be verbally summarized for a firm i as the difference between: 1. The 1% worst-case return of the financial industry, conditional on the return of firm i being at its VaR, and 2. The 1% worst-case return of the financial industry, conditional on the return of firm i being at its median. Mathematically, the implicit definitions,

$$VaR(r^i) : P\{r^i < VaR(r^i)\} = 1\% \tag{23}$$

r^i is the return of firm i. This form allows various arguments. $VaR(r^i)$ can be thought of as the 1% worst case return of firm i, unconditionally. The following two $CoVaR$ measures will be used to define $\Delta CoVaR^i$

$$CoVaR^i_{1\%} : P\{R_I < CoVaR^i_{1\%} | r^i = VaR(r^i)\} = 1\% \tag{24}$$

$$CoVaR^i_{50\%} : P\{R_I < CoVaR^i_{1\%} | r^i = \overline{r^i}\} = 1\% \tag{25}$$

where R_I is the return of the financial industry, $\overline{r^i}$ is the median return of firm i, and what has changed between (24) and (25) is the conditioning ($r^i = VaR(r^i)$ versus $r^i = \overline{r^i}$) rather than the probability (1%). So, in words, $CoVaR^i_{1\%}$ is the 1% worst case return of the financial industry, conditional on the return of firm i being at its VaR, and $CoVaR^i_{50\%}$ is the 1% worst case return of the financial industry, conditional on the return of firm i being at its median. Said another way, these $CoVaR$'s are conditional industry VaR's. Now we can define Adrian and Brunnermeier's proposed measure of firm i's contribution to systemic risk:

$$\Delta CoVaR^i = CoVaR^i_{1\%} - CoVaR^i_{50\%} \tag{26}$$

where $\Delta CoVaR^i$, often just call "CoVaR" in other documents, is the measure of the

firm i's systemic risk. It is the $\Delta CoVaR$ of firm i. This number is usually negative, but often by convention reported as positive, a convention I avoid for clarity. More negative $\Delta CoVaR^i$'s mean more systemic risk.

For starters, consider the robust syste*matic* risk case with no systemic externalities, no assumptions made over the distributions of firm-specific shocks or returns, and only one factor:

$$r_i = b_i\lambda + b_i g + \varepsilon_i$$
$$R_I = b_I\lambda + b_I g + \varepsilon_I$$

Where the second equation is for the industry. Whatever one thinks the difference is between systematic and systemic risk, one does not imagine that this set-up includes anything other than systematic and idiosyncratic risk. According to the original paper regarding what CoVaR measures: "To the extent that it is causal, it captures the risk spillover effects that institution i causes on the industry." The measure should obviously be zero in this clear case of only systematic risk:

$$\Delta CoVaR^i = CoVaR^i_{1\%} - CoVaR^i_{50\%} = 0?$$
$$First,\ \ r_i = b_i\lambda + b_i g + \varepsilon_i \Rightarrow g = \frac{1}{b_i}[r_i - \overline{r_i} - \varepsilon_i]$$
$$R_I = \overline{r_i} + \frac{b_I}{b_i}[r_i - \overline{r_i} - \varepsilon_i] + \varepsilon_I$$

$$\Delta CoVaR^i = CoVaR^i_{1\%}\ -\ CoVaR^i_{50\%}$$
$$\Delta CoVaR^i = VaR[b_I\lambda + \frac{b_I}{b_i} * (VaR(r_i) - \overline{r_i} - \varepsilon_i) + \varepsilon_I]\ -\ VaR(b_I\lambda - \frac{b_I}{b_i}\varepsilon_i + \varepsilon_I)$$

At this point, note that λ, $\overline{r_i}$, and $VaR(r_i)$ are not random variables. Also recognize that while $VaR(\tilde{a}+\tilde{b})$ does not generally equal $VaR(\tilde{a})+VaR(\tilde{b})$, it is true that $VaR(a+\tilde{b}) = a + VaR(\tilde{b})$, where the tilda ($\tilde{}$) over the a or b denotes a random variable. The intuition for this equality is that the entire pdf curve is shifted by adding the constant a, so that the VaR itself is just shifted by a.

$$\Rightarrow \Delta CoVaR^i = [\frac{b_I}{b_i}VaR(r_i) - \frac{b_I}{b_i}\overline{r_i} + VaR(\varepsilon_I - \frac{b_I}{b_i} * \varepsilon_i)]\ -\ [VaR(\varepsilon_I - \frac{b_I}{b_i}\varepsilon_i)]$$
$$= \frac{b_I}{b_i} * [VaR(r_i)\ -\ \overline{r_i}]\ \ < 0$$

Negative $\Delta CoVaR$'s imply positive systemic risk, but should be zero here. Also, ceteris paribus, higher firm beta *lowers* this measure of the systemic risk contribution of the firm.

With the original assumptions and model introduced in Section 4 of the job market paper, Optional assumptions A4O, A5O, and A6O (distributional assumptions) not necessary, we can analyze what happens to CoVaR when there *is* systemic risk from the firm:

$$r_i = b_i\lambda + b_i g + \varepsilon_i$$

$$R_I = \overline{r_i} + b_I g - f(r_i) + \varepsilon_I$$

As the preferred version of CoVaR is defined in direct relation to the returns of the financial industry as a whole (with the idea that the firm is contributing to industry risk), the ISR function above was defined impacting the industry.$g = \frac{1}{b_i}[r_i - \overline{r_i} - \varepsilon_i]$

$$\Delta CoVaR^i = VaR[\overline{r_i} + \frac{b_I}{b_i}(VaR(r_i) - \overline{r_i} - \varepsilon_i) - f(VaR(r_i)) + \varepsilon_I] - VaR[\overline{r_i} - \frac{b_I}{b_i}\varepsilon_i - f(\overline{r_i}) + \varepsilon_I]$$

Now the terms $\frac{b_I}{b_i}VaR(r_i),\ f(VaR(r_i)),\ \frac{b_I}{b_i}\overline{r_i}$, and $f(\overline{r_i})$ are not random variables.

$$\Rightarrow \Delta CoVaR^i = [\overline{r_i} + \frac{b_I}{b_i}VaR(r_i) - \frac{b_I}{b_i}\overline{r_i} - f(VaR(r_i)) + VaR(\varepsilon_I - \frac{b_I}{b_i}\varepsilon_i)]$$

$$- [\overline{r_i} - f(\overline{r_i}) + VaR(\varepsilon_I - \frac{b_I}{b_i}\varepsilon_i)]$$

$$\Rightarrow \Delta CoVaR^i = [\,\frac{b_I}{b_i}(VaR(r_i) - \overline{r_i})\,] + [\,f(VaR(r_i)) \;-\; f(\overline{r_i})\,]$$

The economic intuition behind the second term (and behind CoVaR) is that a negative externality should be large ($f(VaR(r_i))$) when returns are very low and almost non-existent ($f(\overline{r_i})$) otherwise. The second term is math bearing out the intuition behind CoVaR. Other systemic risk measures do not have this aspect. They assume that the non-Gaussian returns come from systematic loading on fat-tailed non-Gaussian factors and on a fat-tailed non-Gaussian idiosyncratic shock. The other point (which is apparently why CoVaR was written as the difference between two conditional VaR's of the *industry* rather than of firm VaR's conditional on the industry) is that the impacted firm (or in this case industry) has the fat tails, while the systemically risky firm does not, raising questions of who we want to regulate.

However, the first term is still present and does not seem to be capturing systemic risk contribution. As mentioned above, the first term depends upon systematic loading when it should not (and increases contribution when loading decreases). CoVaR was an interesting attempt at using economic intuition to determine what function would capture spillovers. Now it is derived and assessed with the benefit of a model that includes spillovers explicitly. Most other measures do not attempt to get at spillovers or systemic externalities. Ultimately, CoVaR does not appear that it will succeed in isolating these spillovers due to 1. The first term decreasing as the systematic risk of the firm (b_i) increases, and 2. I would imagine that $f(\cdot)$ is near zero for most return realizations and only begins to really take-off with extreme, low, rare realizations of r_i, such as in a financial

crisis. Using the 1% $VaR(r_i)$ for the argument of $f(\cdot)$ seems insufficient to capture this[14], or at least insufficient to capture it well enough to make up for problem 1 just listed.

3.6 Part 3 Conclusion

The portfolio optimization model with endogenous-sdf in Part 1 uncovered many new financial and economic intuitions about asset prices, capital allocation, welfare, and systemic risk - as well as bail-outs and their potential for direct effects and allocation distortion effects.

In this Part 3, Sections 3.3, 3.4, and 3.5 leveraged the arbitrage pricing theory model from Section 3.2:

1. to clarify the types of systemic risk - especially systematic versus systemic and interfirm systemic versus inter-industry systemic
2. to establish the basic pricing and arbitrage pricing of assets when ISR is present
3. to show some *implications* of systemic risk for return distributions:
   (a) that idiosyncratic and firm-specific risks (which were formerly synonymous) are not the same thing when systemic risk is present
   (b) that systemic risk changes a firm's returns from Gaussian
   (c) that ISR is naturally decomposed into its three types of components within the returns of the impacted firm or impacted real industry
   (d) that there is positive correlation of a systemic firm's idiosyncratic risk with another firm's idiosyncratic risk
   (e) that the systemically risky firm is not the one that has the distorted return distributions; the firms that it impacts are.
4. to derive CoVaR in terms of the parameters of an asset pricing model to see how CoVaR made a solid intuitive attempt to capture spillovers but ultimately missed, and why.

The factor model has potential to add more to the exciting work being done on systemic risk and returns by deriving further implications about the joint distributions of returns that stem from ISR. So much work in both asset pricing and systemic risk has taken non-Gaussian distributions as given and tried to derive what they imply. In the asset pricing literature, the derivations focus on what the distributions imply for diversification, allocation, and pricing. In the systemic risk literature, the derivations focus on what the distributions imply for systemic risk and the need to regulate.

[14] Acharya et al. (2010) argue that if returns have a component that follows a power law distribution, then 1% $VaR(r_i)$ values can be given a correction coefficient and be sufficient to inform us about what 0.1% or 0.01% $VaR(r_i)$ values would be.

Finally, there are many possible empirical uses. In future work, using a data-smooth or other appropriate econometric technique to back out the form of the externality will yield an interesting paper if the results are compelling about the systemic risk. The results may validate one theory of the mechanism for the systemic externality above the others. It may show a compelling functional form of the systemic risk from the financial sector to the real economy. "Compelling" would be if realized systemic risk is near zero for positive returns of the financial industry and it is smoothly but rapidly increasing as the financial industry's returns drop - combined with a rather low and flat systemic spillover from the real industry to the financial. This could validate the common proxy that researchers often use for systemic crisis: that no crisis and no impacts occur for industry returns above some threshold and financial crisis occurs below that threshold, resulting in significant losses to the real economy. Alternatively, it could call this idea into question.

This could end up being a valuable contribution if it takes hold. The well-known pitfalls of applying Granger causality to measure externalities are absent, pitfalls such as the cyclicality of returns through business cycles appearing as causality, or especially the difficulty of splicing out the consequences of information revelation effects from causal economic impacts. The empirical application of the factor model over time, with a data-smooth for instance, necessarily measures effects. This need not be limited to inter-industry work. Having a way to map out the externality opens up empirical options to test a wide variety of theories that could not be tested in other ways.

Other extensions would be stochastic volatility, how to tell if changes are sliding along the systemic risk curve or represent a shift of the curve, and many others. The start would be a well-researched paper that cleanly delineates what various theories imply about results, how to select functional forms for which to estimate parameters, and other empirical questions.

Finally, empirical work is not limited by tractability. Complex functional forms depending upon aggregate results, or even coalitions, are possible. Work of this type would be somewhat similar to the research on financial networks but without the strong simplifying assumptions, or the preconceived ideas about what is causing the effects. Empirical methods based on A.P.T. factor models with interfirm systemic externalities can agnostically test theories about these externalities, rather than assume a mechanism and measure its parameters. The answers from such studies could go a long way toward "protecting the financial system against the risk spillovers and externalities from systemically important financial institutions... spillovers across institutions that can arise from direct contractual links and heightened counterparty credit risk, or can occur indirectly through price effects and liquidity spirals." Adrian and Brunnermeier (2011).

4 Measuring Systemic Risk

4.1 Introduction

Appropriate regulation of banks and other financial institutions is a long debated topic, and regulations have changed over time. One of the main reasons for regulating such entities is to reduce the risk that their poor performance could impose on the larger economy. Historically, this risk has been difficult to impossible to assess. Can we develop metrics to assess it? And, just as important, can we develop rigorous methods to test and assess those metrics? In this paper, two formerly disjoint strands of literature inform these questions. The first still-developing literature attempts to assess the level of systemic risk within one institution. The second is the analysis of systemic risk in the commercial banking sector before and after Prompt Corrective Action and other concurrent regulatory changes.

Research into systemic risk theory has exploded over the last several years, including attempts to econometrically measure systemic risk for a single firm, summarized in Billio et al. (2010). One of the best known attempts is $\Delta CoVaR$ (Adrian and Brunnermeier (2011))[15], which seeks to use a panel of return data to measure a firm's contribution to systemic risk. The firm's contribution to systemic risk is viewed as a spillover from the firm onto the financial industry - while remaining agnostic to the mechanism causing that spillover. Other systemic risk measures based on panel data have also been offered recently. Some, including $\Delta CoVaR$, are already in use.

Often due to their complexity, very few if any of these measures have been put to the test with any of the standard econometric methods commonly used to achieve clean identification. Having no good tool for evaluation of their systemic risk measure, researchers will often (quite sensibly) rely on other ways to assess their measures, such as emphasizing the intuition, showing strong correlations with things one would expect to be related to systemic risk, and other *ad hoc* methods. We need ways to rigorously test and evaluate such systemic risk measures. In this paper, I use an older strand of literature to accomplish this.

The other strand of literature is the analysis of systemic risk in the U.S. commercial banking sector before and after Prompt Corrective Action and other concurrent regulatory changes in 1993 (hereafter PCA-O). The affects of PCA-O were analyzed and discussed in the literature before, during, and after its roll-out. This paper combines these strands of literature to provide a test of $\Delta CoVaR$ and the effect of extensive regulation on it.

In a difference-in-difference test, the PCA-O regulatory changes are used as a single

[15]Note that although I focus on one particular measure, the difference-in-difference test applies to any newly proposed metric to help gauge its value. Furthermore, because a measure could fail, succeed, or do something in between in a number of ways, much could be learned about what is actually happening when the measure is applied empirically in real-world situations hoping to assess causes and effects. Finally, imposing the need to put the measure into a difference-in-difference framework could, and did in the case of $\Delta CoVaR$, clarify thinking about a measure's components.

treatment affecting commercial banks but not investment banks. The sheer magnitude of the treatment, its exogeneity to firm operations, the periods of regulatory calm pre- and post-treatment, and the similar risk profiles of investment and commercial banks in the United States during the 1990's make this a nearly ideal diff-in-diff scenario that could be used to assess any other systemic risk measure with clean identification. Other researchers may also choose to exploit the the unique circumstances surrounding these regulatory changes for answering other systemic risk questions.

In the case of $\Delta CoVaR$, there are additional benefits to such a study. As $\Delta CoVaR$ has not been built up from any theory, the meaning of any particular value of $\Delta CoVaR$ is unclear. Benchmarking the (causal) impact of an extensive regulatory roll-out on $\Delta CoVaR$ can establish a sense of magnitude for a measure that currently has none. Also, it is not immediately clear what frequency of return data is ideal. When $\Delta CoVaR$ first came out, it used daily returns but changed to weekly. It can be defined on any frequency. Both weekly and daily can be tested and any differences interpreted.

A side benefit of imposing the need to put a systemic risk measure into a difference-in-difference framework could be a clarification of its components. This happens with $\Delta CoVaR$. The difference-in-difference study will highlight a natural decomposition of $\Delta CoVaR$ into its parts - a component driven by fluctuations in the financial industry and a component driven by conditioning.

Finally, As part of this diff-in-diff study, I discuss the various existing $\Delta CoVaR$ measures and develop a new one for assessing the impact of exogenous changes, such as regulation. This new measure also allows computation of standard errors so that we may know how accurately $\Delta CoVaR$'s are being estimated. This is important if it is going to be used as the basis for firm assessments (even more so for regulatory charges or penalties).

The literature survey, Section 2, begins below by discussing the two strands of literature: PCA-O and its effects on the one hand, firm-specific risk measures, especially $\Delta CoVaR$, on the other. Even though three different versions of $\Delta CoVaR$ exist (unconditional, conditional, and forward-looking), I introduce the reader to $\Delta CoVaR$ by reviewing only the standard, unconditional version in Section 2. Section 3 discusses the data, the groups, and the treatment. Section 4 discusses all three versions of $\Delta CoVaR$ and why none is appropriate for time-series assessment of exogenous factors (such as regulation), and develops a new version - unconditional rolling $\Delta CoVaR$. This is the only version of $\Delta CoVaR$ appropriate for testing in a difference-in-difference framework; standard errors are also derived for this version. Section 5 outlines the estimation strategy with the unconditional rolling $\Delta CoVaR$. Section 6 presents the results, which are discussed in Section 7. Section 8 decomposes $\Delta CoVaR$ into its parts, and Section 9 concludes.

4.2 The Two Strands of Literature

4.2.1 PCA and Concurrent Changes

One of the biggest changes in banking regulation in the United States occurred in the early nineties in response to increasing bank failures, the S&L crisis, and mounting international banking integration, see Benston and Kaufman (1997); Berger et al. (1995). This period saw the Basel I Accord, which was applied to the U.S. in the early 1990's with most of the changes occurring in 1993 as part of the PCA implementation section of the FDIC Improvement Act (FDICIA). The law was designed with the goal of reducing systemic banking risk without significant damage to competitiveness, Treasury (1991). All assets and off-balance sheet activities were assigned risk weights according to their perceived risk, and banks were required to hold a certain percentage as capital, as described in Berger and Udell (1994). In addition to being a more subtle requirement, capital requirements were overall more stringent. PCA also implemented changes outlined in the body of FDICIA, including but not limited to more stringent leverage requirements. Another area of change regarded closure policies. Closure policies that banks can expect to face have been shown to affect their risk behavior and resulting credit risk, per Davies and McManus (1991). It seems reasonable to expect an impact on systemic risk as well. PCA also allowed for intervention into banks' operations before closure, known as "structured early intervention and resolution" (SEIR), which also lowered risk, Benston and Kaufman (1997).

Overall, PCA and other concurrent regulatory changes, implemented in 1993, were viewed as the most extensive regulatory attempt to reduce systemic risk in U.S. banks since the Great Depression, Benston and Kaufman (1997); Berger et al. (1995); Aggarwal and Jacques (1998, 2001). They included three significant changes that year: Basel I, PCA, and the new regulatory requirement on unweighted total assets. No other years in this study have such significant regulatory changes. Prior to this, it is widely believed that banks had built up very high levels of systemic risk from the mid 80's until PCA-O went into effect (See in particular Avery and Berger (1991), and many others: Benston and Kaufman (1997); Aggarwal and Jacques (2001), etc.).

PCA-O represents a nearly ideal treatment. It exogenously changed the behavior of firms. It was applied to a clear class of firms (commercial banks) and not to the control group (investment banks). It had a clear goal of impacting risk. It was applied at a point in time - in particular applied over the period of 1993, while the years before and after were free of extensive, from-the-ground-up regulatory change, which was further insured by cutting off the post period in 2000.

More recently, the most up-to-date and sophisticated look at the impact of the law confirms reduced risk. In particular, using the most recent pre-CoVaR techniques, the method of Aggarwal and Jacques (2001) "examines and finds effective" FDICIA and PCA regulation. They further demonstrate that banks were not able to find ways to increase their credit risk to make up for increases in capital ratio. Even the most pessimistic reviews, such as Hovakimian and Kane (2000), accept that it was some improvement. Overall, though vigorous, the only debate remaining is about how large the reduction of industry-wide systemic risk was.

To the extent that researchers of PCA-O distinguished between systemic and systematic risk, which was sometimes little or none, they often seemed to believe that reducing firm risk and/or correlations was sufficient. Many or most of the regulations in PCA-O, however, look quite a bit like regulations being proposed today, making the identified treatment effect of PCA-O particularly relevant.

4.2.2 ΔCoVaR, a Measure of a Firm's Contribution to Systemic Risk

A second, developing literature attempts to assess the level of systemic risk within one institution. We consider here the types of measures that essentially estimate the joint pdf of some financial parameter of the firm and the industry, and derive a metric from that joint pdf. This excludes attempts to use time lags or Granger causality as the essential feature of the core measure. With very rapid trading, or with systemic spillovers that result from asset-price effects, systemic risk spillovers could be nearly instantaneous. The joint pdf is estimated with a panel of returns and is often assumed to be constant over the time frame used to get the estimated joint pdf. Systemic expected shortfall (Acharya et al. (2010)) is one examples of this type of measure. There are others, see Billio et al. (2010); Engle and Manganelli (2004); Huang et al. (2009) for summaries and examples. Systemic expected shortfall is not just the expected shortfall of one of the financial characteristics of one firm (in this case capitalization); it is expected shortfall whenever the financial industry is under-capitalized. This obviously requires a joint pdf of the firm with the industry rather than just the firm's pdf.

The most widely cited of these types of measures has been $\Delta CoVaR$ (Adrian and Brunnermeier (2011)). That original paper (Adrian and Brunnermeier (2011)) proposes three different $\Delta CoVaR$'s (unconditional, conditional, and forward-looking). In this section, I review the standard, unconditional $\Delta CoVaR$. Section 3 discusses all three versions and why none is appropriate for testing in a diff-in-diff framework, and develops a new version - unconditional rolling $\Delta CoVaR$ - which is appropriate[16].

$\Delta CoVaR$ stands for Co-Value-at-Risk, and can be verbally summarized for a firm i as the difference between: 1. The 1% worst case return of the financial industry, conditional on the return of firm i being at its VaR, and 2. The 1% worst case return of the financial industry, conditional on the return of firm i being at its median. Mathematically, the implicit definitions,

$$VaR^i : P\{r^i < VaR^i\} = 1\% \tag{27}$$

r^i is the return of firm i. VaR^i can be thought of as the 1% worst case return of firm i, unconditionally. The following two $CoVaR$ measures will be used to define $\Delta CoVaR^i$

$$CoVaR^i_{1\%} : P\{R < CoVaR^i_{1\%} | r^i = VaR^i\} = 1\% \tag{28}$$

[16]Unconditional rolling $\Delta CoVaR$ is most reflective of Adrian and Brunnermeier's core measure, unconditional $\Delta CoVaR$, to give $\Delta CoVaR$ its best chance in a difference-in-difference framework.

$$CoVaR^i_{50\%} : P\{R < CoVaR^i_{1\%} | r^i = \overline{r^i}\} = 1\% \tag{29}$$

where R is the return of the financial industry, and $\overline{r^i}$ is the median return of firm i. So, in words, $CoVaR^i_{1\%}$ is the 1% worst case return of the financial industry, conditional on the return of firm i being at its VaR, and $CoVaR^i_{50\%}$ is the 1% worst case return of the financial industry, conditional on the return of firm i being at its median. Said another way, these $CoVaR$'s are conditional industry VaR's. Now we can define Adrian and Brunnermeier's proposed measure of firm i's contribution to systemic risk:

$$\Delta CoVaR^i_{1\%} = CoVaR^i_{1\%} - CoVaR^i_{50\%} \tag{30}$$

where $\Delta CoVaR^i_{1\%}$, often just call "CoVaR" in other documents, is the measure of the firm i's systemic risk. This number is usually negative, but often by convention reported as positive, a convention I follow by defining:

$$\Delta CoVaR^i \equiv -\Delta CoVaR^i_{1\%} = CoVaR^i_{50\%} - CoVaR^i_{1\%} \tag{31}$$

(5) is the $\Delta CoVaR$ of firm i. Higher $\Delta CoVaR^i$'s mean more systemic risk.

In time series, Adrian and Brunnermeier document correlation between unconditional firm $\Delta CoVaR$ and certain firm characteristics which are theoretical determinants of systemic risk. This is descriptive information about when firms have higher $\Delta CoVaR$'s. For example, consider the fact that high leverage is correlated with high $\Delta CoVaR$. This may reflect high $\Delta CoVaR$ causing firms to respond with high leverage, the converse, or the fact that an unobserved factor causes both high leverage and high $\Delta CoVaR$. However, the magnitudes of these correlations cannot be used to give any indication of what we should consider low or high values of $\Delta CoVaR$.

By how much does extensive regulation reduce a firm's systemic risk? How well does $\Delta CoVaR$ capture systemic risk? Another question is about what values of $\Delta CoVaR$ represent little or much systemic risk: though there is an intuition for $\Delta CoVaR$ capturing systemic risk, and $\Delta CoVaR$ fulfills many of the properties one would like to see in such a measure (directionality and the cloning property for example), the lack of a theoretical derivation for the metric yields questions about what should be considered a lot of systemic risk.

In Section 3, I determine the change in $\Delta CoVaR$ that can be attributed to PCA-O to inform the questions above. To whatever extent researchers have opinions about how high systemic risk was in the period leading up to PCA-O and how great the reduction from PCA-O, isolating its causal impact on $\Delta CoVaR$ informs that question.

4.3 Data, the Two Groups, and the Treatment

The database comes from COMPUSTAT with an embedded permanent company identifier (permno) from the CRSP database. Using a dummy variable indicating quarter, this was merged with a weekly CRSP data file in which I calculated market cap and leverage. Both data files, and hence the merged file, were for the period from January of 1986 to December

of 2000. This merged database is used as the source for analysis. Daily returns of common stock are scaled by leverage to give daily or weekly returns of company assets. The more accurate SIC codes from COMPUSTAT were used to select companies while those from CRSP were ignored. All of these choices (except the time period) follow the creators of $\Delta CoVaR$, Adrian and Brunnermeier (2011). The definition of the financial industry and data-processing of the industry also follow the original.

As investment banks were not subject to PCA-O, they are the control group. True investment banks not part of bank holding companies numbered only 64 during the period of study. The treated group of commercial banks is larger, numbering more than 800. The analysis is limited to investment banks of SIC code 6211 and commercial banks of SIC code 6020. The largest of the 6020's were Bank of New York Mellow Corp., Wachovia Corp., Wells Fargo & Co., Bank of America Corp. The largest of the 6211's were Goldman Sachs Group Inc., Morgan Stanley, and Merril Lynch.

On the key characteristics used in the calculations is A, market capitalization times leverage. Leverage is defined as book assets over book equity. It is one of the most commonly identified culprits of excessive risk and a favorite target of regulation. Comparing the groups:

Item (mean)	Investment Banks	Commercial Banks	Ratio (Inv./Comm.)
Assets (book)	\$17.4B	11.30B	1.54
Market Cap	1.50B	1.39B	1.09
Leverage	9.7	13.1	0.74
A = (Mcap*Lev)	34.3B	19.2B	1.79

One thing that can confound difference-in-difference tests are time-dependent, group-dependent shocks. A full description of commercial and investment banks in the 1990's is beyond the scope of this paper. But, in sum, commercial banks take deposits and make private and commercial loans, sometimes selling them. Commercial banks make money on deposits, loans, and off-balance sheet activities. Investment banks do not take deposits but underwrite securities. The bases of underwritten securities include private or commercial debt, or equities. Investment banks make most of their money from underwriting, although many make a lot from proprietary trading and brokerage as well.

Considering risk, both groups would be impacted similarly through their two most prominent channels: lending and fluctuations in holdings for commercial banks, and underwriting and fluctuations in holdings for investment banks. In the former case, there is exposure is to the market for investments that are funded with debt[17], the market for consumption funded with debt, and asset prices. In the latter case, there is exposure is to the market for investments funded with debt[18], the market for investments funded with equity, and asset prices. The groups are clearly subject to similar risks. In particular,

[17] as bank loans

[18] as bonds or commercial paper

cycles in the macro-economy in general, and in credit markets in particular, would not appear, at least *a priori*, to give rise to time-dependent, group-dependent shocks during the study period.

However, with a difference-in-difference estimation strategy, it is not necessary to motivate *ex ante* that the dependent variable is not hit with significant time-dependent, group-dependent shocks, other than the treatment. Nor is it sufficient. When presenting results, the dependent variable for the control and treatment groups is plotted over time and compared, looking for evidence of such shocks year-to-year. This method would not, of course, uncover any shocks that affected one of the groups beginning exactly during the treatment year and lasting until the end of the treatment period. It may also not uncover shocks that evolved slowly far from the treatment. So selection of the treatment year, the control period, and the treatment period are important.

Basel I and some of the provisions that ultimately ended-up in PCA were finalized (though not implemented) in the late 1980's. Though there were important macro-economic events (the credit crunch in the early 90's, the economic recovery beginning in 1992), trends that could conceivably confound the difference-in-difference estimation strategy (time-dependent, group-dependent shocks) do not come to mind for the period from 1986 through 2000. This is not true of the period after 2000, where factors expected to have differential impacts arise. The key difference in these time periods is that the business models of the groups remained stable in the 1986 through 2000 time period.

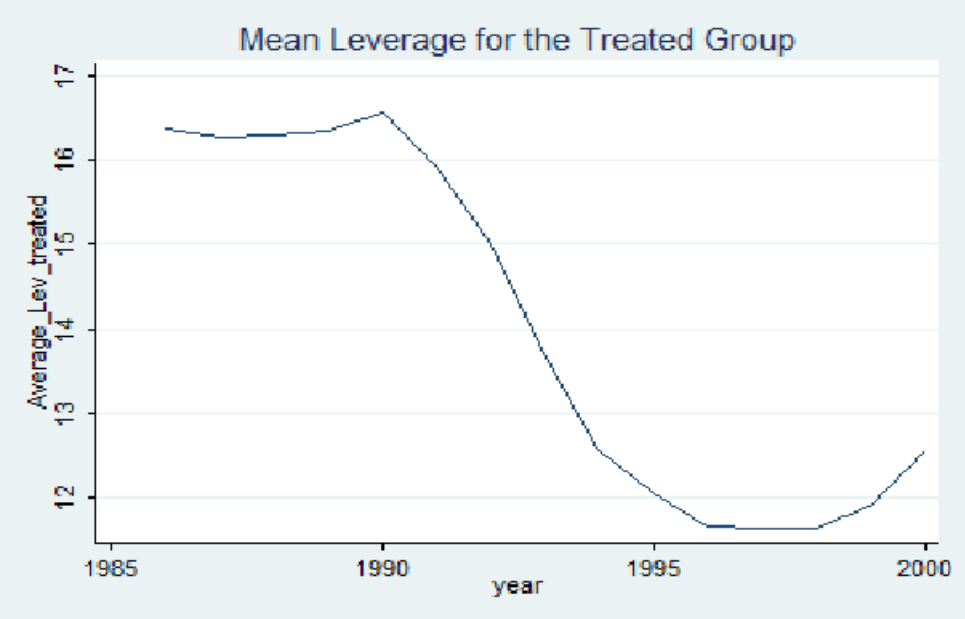
Massive securitization and shadow banking became significant in 2006, and mushroomed in 2007. This alone makes these years inappropriate for including in the analysis. The financial crisis itself also creates unworkable challenges. One is the resulting regulatory flux. An even bigger issue is the fact that banks began converting into bank holding companies - exiting from the treatment group to receive TARP money. Because of these issues, 2005 would be the last year that could reasonably be used in this study.

I also eliminate the period 2001-2005. These are the years most likely to be unrepresentative for three reasons. First, the Commodity Futures Modernization Act of 2000 was significant and should have different impacts on the behavior of investment banks versus commercial banks. Such impacts would represent the kind of time-dependent, group-dependent shock that can cause problems with a difference-in-difference estimation strategy. Eliminating periods after 2000 obviates the need to assess differential effects of this law. Secondly, the repeal of GlassSteagall was passed in 1999 and went into effect in 2000. Thirdly, these years are furthest from the treatment, allowing more time for the evolution of differences in the groups which are not in the pre-treatment period. Finally, this cut-off leaves a symmetric seven-year pre-treatment period and a seven-year post-treatment period, both extremely calm periods in banking regulation change.

The bulk of the PCA-O changes occurred over the calendar year of 1993, Benston and Kaufman (1997); Aggarwal and Jacques (1998). The changes that did not occur in 1993 are smaller but do represent an imperfection in this study. In addition they do not appear to be stacked more heavily into pre-treatment or post-treatment period - for example into 1992 as opposed to 1994, or *vice versa*. Finally, the magnitude of the treatment and the viewing of trajectories mitigate these issues. Accordingly, commercial banks have a

treatment status of 0 from 1987 to 1992 and 1 from 1994 to 2000.

It is worth emphasizing that the treatment is the collection of stringent regulatory changes implemented throughout 1993, rather than direct exogenous manipulations of firm characteristics. Because of the magnitude of the treatment and the clean identification, this scenario informs us about what might be achieved with these types of direct regulations. The treatment is net of firm response. The firms changed in observable and unobservable ways to comply with all the regulations. As one example, the impact of the regulation on the key firm characteristic of leverage:



4.4 Estimating the Dependent Variable

4.4.1 Which ΔCoVaR?

In this analysis, the dependent variable, $\Delta CoVaR$, must be estimated. Adrian and Brunnermeier (2011) provide three $\Delta CoVaR$ measures: conditional, unconditional, and forward-looking. Which, if any of these measures, is appropriate for assessing the impact of an exogenous shock in a diff-in-diff framework?

Unconditional ΔCoVaR?

Being time-invariant, the unconditional $\Delta CoVaR$ obviously cannot work here. A treatment-effect estimate *could* be obtained by estimating two unconditional $\Delta CoVaR$'s for each firm, one for the pre-treatment period and one for the post-treatment period. This would

facilitate estimation but not determination of whether the groups have parallel trajectories, an important diff-in-diff robustness check. It would also obviate any ability to see if the treatment effect is focused on the period near the treatment. Thus, unconditional $\Delta CoVaR$ is not appropriate for the current study.

Conditional ΔCoVaR?

The essence of this sub-section is to point out that all time-variation in a firm's conditional $\Delta CoVaR$ comes from variation in macroeconomic factors that are external to the firm. Hence, it cannot be used to assess changes to $\Delta CoVaR$ that come from the firm changing in the face of new regulation, and it cannot be used for this study.

This is not made explicit in the original paper. To see this, note that equation 9 at the top of page 15 of the original paper can be solved to give:

$$\Delta CoVaR^i_{.01,\,t} = \hat{\beta}_{.01,\,system}(\hat{\alpha}^i_{.01} - \hat{\alpha}^i_{0.5} + \hat{\gamma}^i_{.01} M_{t-1} - \hat{\gamma}^i_{0.5} M_{t-1}) \tag{32}$$

So that, after the regression is run to obtain the coefficients, all time variation comes from the independent variables M_{t-1}. Examples of the components of M include VIX, T-bill rates, and credit spreads.

Thus, conditional $\Delta CoVaR$ measure is not appropriate for econometric identification because it assumes an underlying model where firm $\Delta CoVaR$ is only a function of the environment and not of the firm characteristics[19], exogenously altered here. This implies that any change in the $\Delta CoVaR$'s of the treated firms which results from the regulatory changes will be assigned to both the pre- and post-treatment periods via the firm's altered factor model - artificially eliminating any treatment effect, making conditional $\Delta CoVaR$ also unworkable here.

Forward ΔCoVaR?

Finally, using forward $\Delta CoVaR$'s for estimation is unworkable because forward $\Delta CoVaR$ predictions use the firm characteristic which are part of the causal chain of the treatment's impact. Secondly, since this difference-in-difference scenario has hindsight, why find the effect of the treatment on predictions of $\Delta CoVaR$'s instead of on $\Delta CoVaR$'s? Thirdly, it would add layers of unnecessary complexity as forward $\Delta CoVaR$'s are themselves predictors of conditional $\Delta CoVaR$. In fact, forward $\Delta CoVaR$'s use firm characteristics, such as leverage, etc., without respect to whether they are changed exogenously, which may be appropriate to Adrian and Brunnermeier's method. However, in this context, it is not clear what it would even mean to find a certain estimated treatment effect of exogenous regulatory changes on forward $\Delta CoVaR$'s (i.e., on predictions of future $\Delta CoVaR$'s based on regressions on firm characteristics). This eliminates the third and final existing measure of $\Delta CoVaR$ for use in an identified difference-in-difference estimation.

[19] More specifically, it is not a function of firm characteristics in time series, which we require. It is a function of the specific firm in the sense that the factor model was obtained for the specific firm from its returns, where $\hat{\alpha}^i_q$, $\hat{\gamma}^i_q$ are obtained by regressing firm returns on the factors.

Unconditional Rolling ΔCoVaR

To solve this problem, I calculate unconditional $\Delta CoVaR$ for one-year windows. This lets us test the essence of Adrian and Brunnermeier's core measure, unconditional $\Delta CoVaR$. Because the estimation comes from just one year of data, estimation error is larger than for the normal unconditional $\Delta CoVaR$. The additional error created in the dependent variable is acceptable because we are comparing groups and so have more data. This allows the required time variation, without the inability of $\Delta CoVaR$ to change with the time-specific treatment as with the conditional $\Delta CoVaR$, or the unclear and convoluted interpretation of the treatment effect as with the forward $\Delta CoVaR$. This gives $\Delta CoVaR$ its best and cleanest chance. The reason for using disjoint time periods is to avoid reusing data and creating hysteresis problems with estimation. However, one could use a longer, rolling window with a trailing-off kernel weighting-function to estimate $\Delta CoVaR$ at each point for other applications to give the best estimate of $\Delta CoVaR$ at a point in time. This would be important for assessing efforts to change the systemic risk of the institution by exogenously changing firm characteristics, because presumably such efforts would act by changing the functional form of conditional $\Delta CoVaR$ as a function of factors, rather than changing the evolution of the factors themselves. Using the same functional form (equation 9 above) would not pick-up the regulator's efforts. The unconditional rolling $\Delta CoVaR$ with weighting solves this problem.

4.4.2 Estimating Unconditional Rolling ΔCoVaR's

Estimates of unconditional rolling $\Delta CoVaR$ are the left-hand-side variable for the diff-in-diff test. The methodology is similar to that used by Adrian and Brunnermeier to estimate unconditional $\Delta CoVaR$'s. First, daily (or weekly, assume daily for this discussion) median and VaR of returns is determined for each firm/year combination where i is the firm with return r and j is the year:

$$VaR^{i,j} : P\{r^{i,j} < VaR^{i,j}\} = 1\%$$

The 1% quantile regression of daily returns on a constant estimates the above $VaR^{i,j}$.

The $\Delta CoVaR$ estimation method is to run a quantile regression against industry returns as applied in the original paper. The definition of unconditional rolling $\Delta CoVaR$ is similar to unconditional $\Delta CoVaR$:

$$CoVaR^{i,j}_{1\%} : P\{R < CoVaR^{i,j}_{1\%} | r^{i,j} = VaR^{i,j}\} = 1\% \tag{33}$$

$$CoVaR^{i,j}_{50\%} : P\{R < CoVaR^{i,j}_{50\%} | r^{i,j} = \overline{r^{i,j}}\} = 1\% \tag{34}$$

$$\Delta CoVaR^{i,j} = CoVaR^{i,j}_{50\%} - CoVaR^{i,j}_{1\%} \tag{35}$$

where R is the daily return of the financial industry; $\overline{r^{i,j}}$ is the median daily return of the firm i during year j; $r^{i,j}$ is the random variable representing the daily return of

the firm i during year j; and $VaR^{i,j}$, defined above, is the VaR of daily return for firm i during year j.

The next requirement is to obtain the quantile regression coefficients for the particular firm *and the particular year*, estimating the coefficients of the following affine quantile model, where t is the day or week:

$$R^t_{.01} = \gamma^{i,j}_{.01} + \beta^{i,j}_{.01} r^{i,j}_k + \varepsilon_{i,j,t} \tag{36}$$

The $\hat{\beta}^{i,j}_{.01}$ so determined is saved for the firm and year from which it was estimated.

Noting that

$$CoVaR^{i,j} = \gamma^{i,j}_{.01} + \beta^{i,j}_{.01} VaR^{i,j} + \varepsilon$$

and

$$CoVaR^{i,j}_{50\%} = \gamma^{i,j}_{.01} + \beta^{i,j}_{.01} \overline{r^{i,j}} + \varepsilon$$

the estimation of (9) for firm i in year j is:

$$\Delta \hat{CoVaR}^{i,j} = \hat{\beta}^{i,j}_{.01} (\overline{r^{i,j}} - \hat{VaR}^{i,j}) \tag{37}$$

In summary, to estimate (11) for one firm in one year[20]: 1). A quantile regression of the firm's daily returns on a constant gives $\hat{VaR}^{i,j}$, 2). A quantile regression of the daily returns of the industry on the firm's daily returns gives $\hat{\beta}^{i,j}_{.01}$, and 3). The sample median $\overline{r^{i,j}}$ estimates the population median.

Unlike the original paper, I estimate and report the confidence interval of the unconditional rolling $\Delta CoVaR$ estimates. This is a C.I. *estimate* because it is calculated assuming a constant joint pdf for the time window. $\Delta \hat{CoVaR}^{i,j}$ confidence intervals come from the standard errors on $\hat{\beta}_{.01}$ and $\hat{VaR}^{i,j}$, plus the assumption that these two estimates are independent:

$$\sigma^2_{\Delta CoVaR} = \hat{\beta}^2_{.01} \sigma^2_{VaR} + (\hat{VaR}^{i,j} - \overline{r^{i,j}})^2 \sigma^2_\beta + \sigma^2_{VaR} \sigma^2_\beta \tag{38}$$

The estimated median $\overline{r^{i,j}}$ is assumed to be a constant, equal to the population median. This assumption is at least fairly innocuous, but probably conservative. This is because $\hat{VaR}^{i,j}$ and $\overline{r^{i,j}}$ are both estimated from the time series of returns of the firm, and a preponderance of high (low) return realizations would simultaneously increase (decrease) both estimates. The resulting positive correlation of the two variables most likely makes the variance of $(\hat{VaR}^{i,j}_{.01} - \overline{r^{i,j}})$ *lower* than σ^2_{VaR}, which is used in the above. Standard errors calculated this way are reported with the results.

[20]The process takes up to a week to run, with "2)" being the most time-consuming. Quantile regressions minimize absolute, rather than squared, deviation. They are solved iteratively and require more calculations.

4.5 Treatment Effect Estimate

The treatment effect is estimated from

$$\Delta \hat{CoVaR}^{i,j}_{.01} = a_i + \delta_j + \tau D_{i,j} + \varepsilon_{i,j} \tag{39}$$

where a_i is a firm fixed effect, δ_j is year fixed effect, and $\widehat{D_{i,j}}$ is the PCA-O treatment dummy. The regression-predicted dependent variable is $\Delta \hat{CoVaR}^{i,j}_{.01}$ $(= \hat{a}_i + \hat{\delta}_j + \hat{\tau} D_{i,j})$. The lack of additional covariates is due to the extensive nature of the regulation, leaving no firm characteristic predetermined.

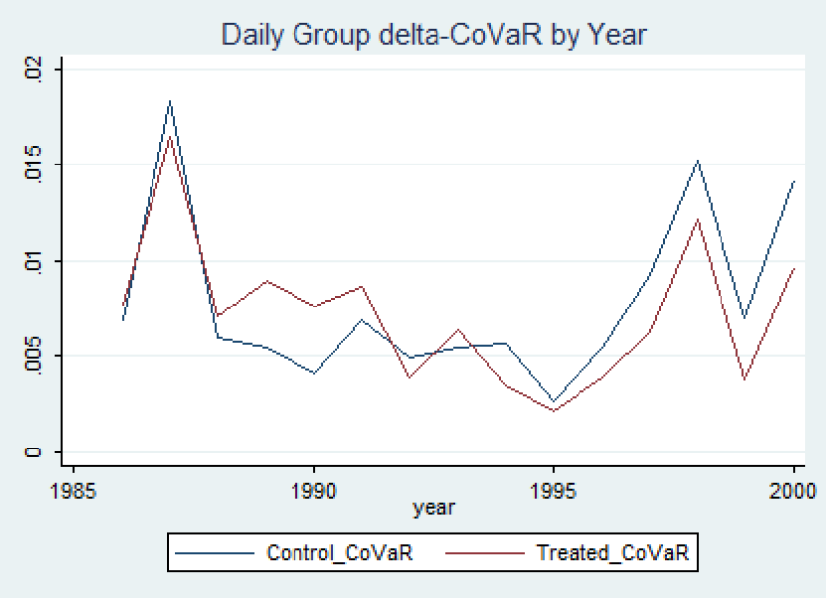
The independent variable is simply whether or not the firm is regulated under PCA-O; measurement error could only come from incorrectly specified SIC's. Furthermore, minor attenuation bias is the only result of even moderate measurement error (in this case mis-assignment).

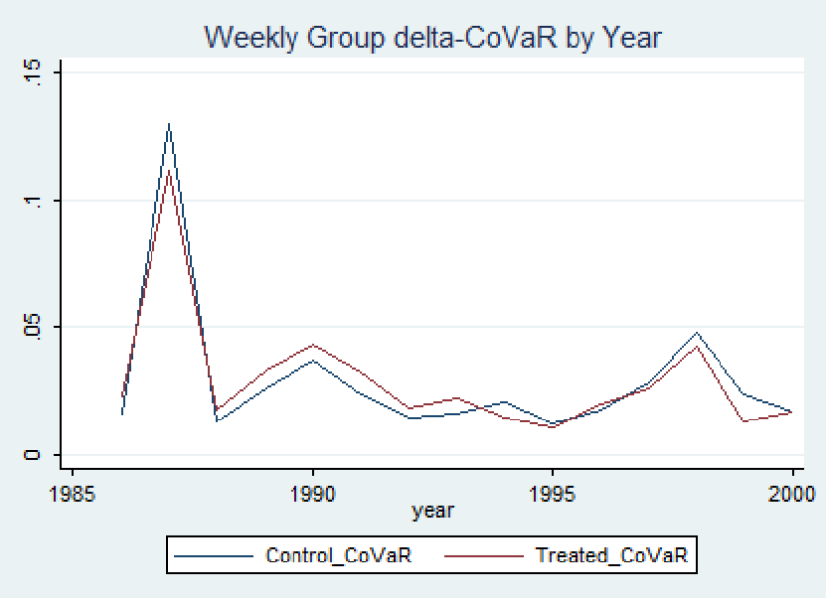
In contrast, the dependent variable will have very high measurement error. However, this is only a concern in that it may wash out the effects. Unless there is very systematic measurement error (correlated with the treatment), there will not be bias in the estimates. The measurement errors in the estimation of the dependent variable get rolled into the $\varepsilon_{i,j}$'s, increasing the standard errors. The potential that $\Delta CoVaR$ estimation errors are correlated within firms for various years necessitates use of robust standard errors.

The precision of the treatment effect estimate could be increased slightly by weighting the regression by the inverse of the dependent variable's standard errors, but the downsides of this option outweigh this benefit. Weighting confuses interpretation of what effect is being estimated and makes the estimates too sensitive to set-up. In addition, the results in the next section do not require further refinement of estimates. Finally, the standard errors of the CoVaR's (of $\Delta \hat{CoVaR}^{i,j}_{.01}$) average 0.165% as calculated using (8), equaling only ~10% of the variance of the CoVaR's. This all but eliminates any gains from weighting:

Variable	Obs.	Mean	Std Dev.	Min.	Max.
S.E. of CoVaR Estimate	1456	.0016486	.0036227	1.40e-06	.0597138

4.6 Results





The figured do not show significant group-dependent, time-dependent shocks as the two groups follow similar trajectories. Also absent from the figure is any discontinuity (or even gradual separation) between groups at the boundary between the pre- and post-treatment periods, 1993. The treatment, PCA-O is regulatory changes that began applying to commercial banks anytime during 1993. The treatment period is 1994-2000, while 1987-1992 is pre-treatment. Overall, no treatment effect is visible.

The only reasonable confounding factor to this type of diff-in-diff study is a time-

dependent, group-dependent shock in 1993. The potential for this is discussed in Section 3. However, because we observe no treatment effect, it would be quite a knife-edge case if there was a treatment effect appearing in 1993 that was hidden by an equal and opposite time-dependent, group-dependent shock appearing only in 1993.

The observations per group varied:

Group variable: firm.

Minimum	average	maximum
1	5.8	15

Number of groups for daily estimation = 896. Total Observations: 5227. Controls Observations: 693.

As individual firms are not tracked throughout the sample, one concern in the database is survivorship bias - the weak firms subject to more stringent regulatory requirements may have trouble surviving the change. This will not be true about the control group. Although I cannot rule out the possibility of this creating minor bias, I emphasize that, at least, there is no *a priori* reason to imagine that firm strength is negatively or positively correlated with $\Delta CoVaR$[21]

The estimation is as follows:

DAILY $\Delta CoVaR$:

Coefficient	Estimate	Std. Err.	t	95% Conf. Interval
Treatment	.00169	.000879	1.92	-.00003 .00341
Constant	.00383	.0066	5.81	.00254 .00512

WEEKLY $\Delta CoVaR$:

Coefficient	Estimate	Std. Err.	t	95% Conf. Interval
Treatment	-.00184	.00214	-0.86	-.00603 .00234
Constant	.0190	.00239	7.95	.0143 .0237

The null cannot be rejected and the confidence interval contains zero in both cases. In the daily case, the point estimate of the treatment represents an *increase* in systemic risk. We do not have statistical significance.

Because $\Delta CoVaR$ is a new measure and is not rigorously tied to any theory, it is not immediately apparent whether the above confidence intervals imply anything about economic significance. In the absence of other guidance, I turn to the mean and standard deviation of all the $\Delta CoVaR$'s in this study to provide some sense of economic magnitude. This gives us some sense of the change in systemic risk, relative to the total systemic risk that existed in the population.

[21]The authors of the original paper do not suggest any evidence of this in the cross-section, which is the relevant relationship for survivorship bias.

DAILY $\Delta CoVaR$:

Variable	Mean	Standard Dev.	Minimum	Maximum
$\Delta CoVaR$	.00686	.0105	-.0603	.0720

WEEKLY $\Delta CoVaR$:

Variable	Mean	Standard Dev.	Minimum	Maximum
$\Delta CoVaR$	.0260	.0384	-.110	.306

The daily point estimate is positive. The weekly point estimate is less than one tenth of the mean of the weekly $\Delta CoVaR$'s and one twentieth of the standard deviation. Within both confidence intervals, the most beneficial treatment effect is the left-hand side of the weekly confidence interval, -0.00603, less than a fourth of the mean and a sixth of a standard deviation.

4.7 Results Discussion

The apparent lack of economic significance from such extensive treatment is surprising, and arguably makes a null result *more* interesting in this case. The result challenges the commonly held beliefs that $\Delta CoVaR$ measures systemic risk and that the regulatory changes of 1993 eliminated most of the systemic risk, or at least greatly reduced systemic risk relative to the amount of systemic risk that existed. In fact, one of these two must be false. At first blush, this is quite a surprise. Perhaps more importantly, most of the changes wrought by PCA-O are similar to many changes being proffered now for reducing systemic risk.

The result also appears to conflict with (or at least require reconciliation *vis a vis*) the results of Adrian and Brunnermeier that improvements in capital ratio or leverage ratio lower $\Delta CoVaR$ in both the cross-section and time series. If these are two key components of PCA-O, and PCA-O added other factors that are widely believed to lower systemic risk further, what explains the discrepancy? I attribute the difference between their cross-sectional result and the results here to the fact that they were comparing a wide variety of companies (the entire financial industry) without isolation of causality. Firms with fundamentally different business models and structural make-up were compared over a 20 year period with one $\Delta CoVaR$ measure to capture the entire period. In short, their cross-sectional result was simply the discovery of correlations.

The discrepancy in their time series result and my result can of course be explained with a similar argument. Perhaps in time series there are unobservable firm characteristics in conjunction with both $\Delta CoVaR$ within a firm on the one hand, and capital ratio and leverage ratio within a firm on the other. However, the full reconciliation of the time series results of Adrian and Brunnermeier are beyond the scope of this paper. In estimating time varying $\Delta CoVaR$, the conditional distribution is assumed to be a function of state variables (i.e. "factors"), and then $\Delta CoVaR$ is calculated directly from this conditional distribution. This means that the $\Delta CoVaR$ for a firm is a function of state variables. The authors hold the function itself constant. All variation through time of the $\Delta CoVaR$ of a

firm comes from variation in state variables only (equation 6 above), not from any changes to the firm's sensitivity to these factors. So, if a characteristic of the firm (leverage for instance) changes through time, and the conditional $\Delta CoVaR$ of the firm also changes through time, the only mechanism by which these two items can be correlated is via correlation between the firm characteristic and state variables - making any relationship between firm characteristics and firm $\Delta CoVaR$ more complicated.

For example, risky bond credit spreads and treasury liquidity-risk spreads are two of the factors. If the economy's credit spread and/or liquidity spread is, on average, higher in years when financial firms have higher leverage, then a time series correlation between a firm's leverage and it's $\Delta CoVaR$ will obtain, even though *the underlying function which determines* $\Delta CoVaR$ *from the factors has not changed*. The meaning of this is more difficult to interpret than the results presented here.

None of this is to say that increasing conditional $\Delta CoVaR$ levels do not predict systemic crises, only that they can only do so by changing factors. Average conditional $\Delta CoVaR$ may still be such a predictor, but the results here show that neither it, nor unconditional $\Delta CoVaR$, are impacted by an exogenous regulatory shock.

4.8 ΔCoVaR Decomposition

One fact, which is not new but was highlighted by calculating rolling unconditional $\Delta CoVaR$'s, is that $\Delta CoVaR$ is a difference in two **conditional** VaR's of the financial industry. It is natural to ask whether this difference scales with the unconditional VaR of the financial industry or not. To clarify notation, define VaR^I as the unconditional VaR of the financial industry and rewrite (2) and (3) as conditional VaR's:

$$\{VaR^I | r^i = VaR^i\} \equiv CoVaR^i_{1\%} \tag{40}$$

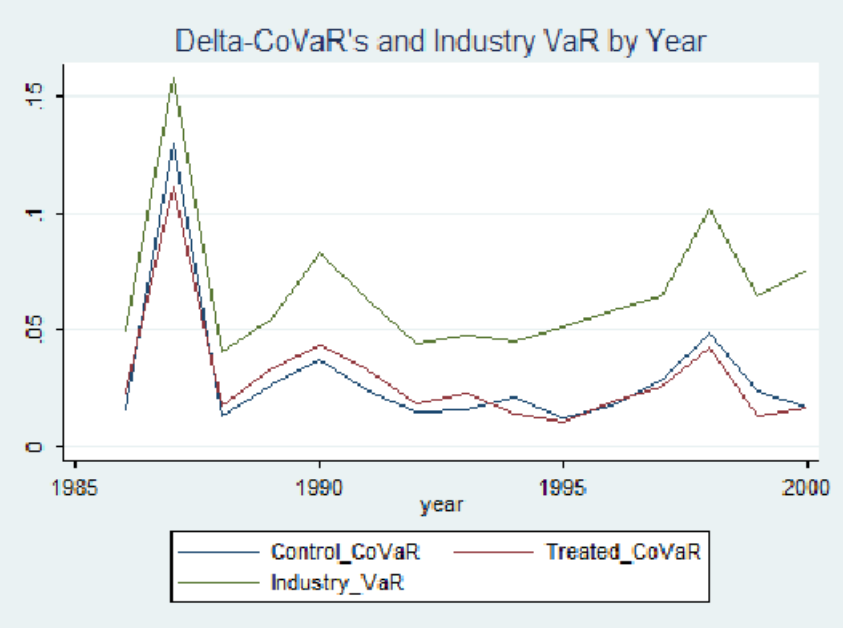
$$\{VaR^I | r^i = \overline{r^i}\} \equiv CoVaR^i_{50\%} \tag{41}$$

$\Delta CoVaR$ could be thought of as the magnitude of the unconditional VaR of the financial industry times the percent change that conditioning on two different states of the firm causes ($\%change$), or simply:

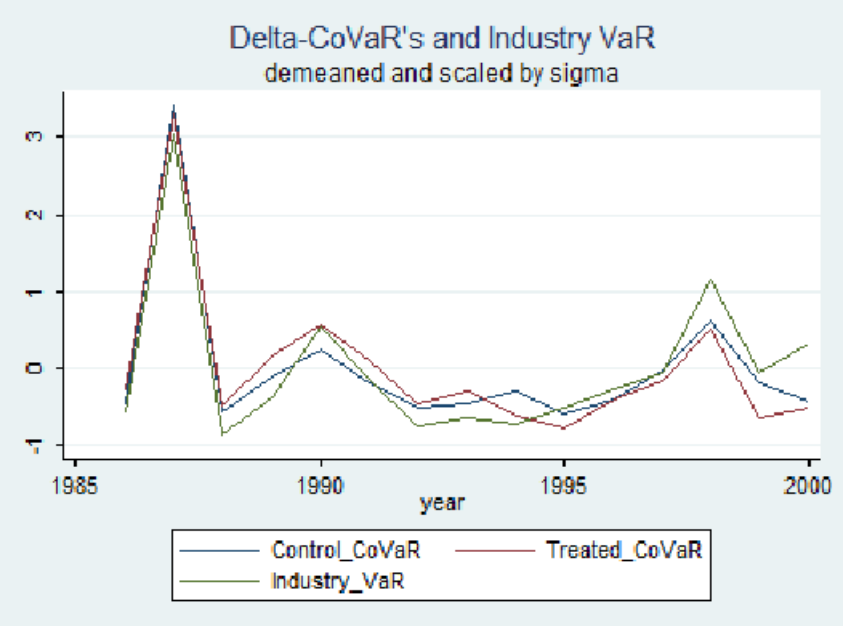
$$\%change \equiv [\,\{VaR^I | r^i = \overline{r^i}\} - \{VaR^I | r^i = VaR^i\}\,] \,/\, VaR^I \tag{42}$$

$$\Rightarrow \Delta CoVaR^i = \%change * VaR^I \tag{43}$$

If variation in the latter factor dominates, then a firm's $\Delta CoVaR$ is caused mostly by differences in conditioning and could, at least conceivably, also be *caused* by how the firm is interacting with the industry, including its spillovers. If variation in the former factor dominates, then the conditioning does not vary and $\Delta CoVaR$ is being driven by factors external to the firm, namely the VaR of the financial industry. At least in the cases of the control and treatment groups, financial industry VaR provides most of the variation. Changes in conditioning are very small:



Looking only at trends:



4.9 Part 4 Conclusion

The various versions of $\Delta CoVaR$ utilized by Adrian and Brunnermeier for regulatory assessment are not ideal for econometric testing of the impact of firm characteristics on

firm $\Delta CoVaR$. Such studies can employ the unconditional rolling $\Delta CoVaR$ developed herein. Such a measure could also be used to assess the progress of a firm in lowering its $\Delta CoVaR$. A method for determining its confidence interval is also developed.

However, the first clean test of causality between any characteristic of a firm (in this case whether it was regulated under PCA-O) and its $\Delta CoVaR$ shows no effect. Maintaining the hypothesis that these regulations lowered systemic risk, this implies that $\Delta CoVaR$ is not capturing systemic risk. In the treatment and control groups, ~85% of the variation in group $\Delta CoVaR$ is explained by variation in the financial industry's unconditional VaR, providing some evidence that variation in a firm's $\Delta CoVaR$ is not driven by its own systemic risk. The circumstances of the extensive and coincident regulatory changes in the United States in 1993 could also be used to assess other measures of systemic risk or other dependent variables believed to be affected by regulation. The most glaring need for future research is to develop a theoretical framework for assessing what $\Delta CoVaR$ and other systemic risk measures capture.

The circumstances in U.S. banking represent a generally unexploited natural experiment that can be used to address questions of risk, banking behavior, regulation, and other questions. The sheer magnitude of the treatment (Prompt Corrective Action and implementation of Basel I) hitting commercial but not investment banks almost entirely in one year, with periods of relative regulatory calm before and after , should be used to inform other important systemic risk and banking questions.

References

Acharya, Viral V., Lasse H. Pedersen, Thomas Philippon, and Matthew P. Richardson, 2010, Measuring systemic risk, *FRB of Cleveland Working Paper No. 10-02* .

Adrian, Tobias, and Markus K. Brunnermeier, 2011, Covar, *FRB of New York Staff Report No. 348* 2008.

Aggarwal, Raj, and Kevin Jacques, 1998, Assessing the impact of prompt corrective action on bank capital and risk, *Economic Policy Review* 4.

Aggarwal, Raj, and Kevin T Jacques, 2001, The impact of fdicia and prompt corrective action on bank capital and risk: Estimates using a simultaneous equations model, *Journal of Banking & Finance* 25, 1139–1160.

Allen, Franklin, and Douglas M. Gale, 2000, Financial contagion, *Journal of Political Economy* 108, 1.

Allen, Linda, Turan G Bali, and Yi Tang, 2012, Does systemic risk in the financial sector predict future economic downturns?, *Review of Financial Studies* 25, 3000–3036.

Andersen, Torben G, and Tim Bollerslev, 1997, Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns, *The Journal of Finance* 52, 975–1005.

Angelini, P., G . Maresca, and D. Russo, 1996, Systemic risk in the netting system, *Journal of Banking and Finance* 20, 853–868.

Arrow, Kenneth J, and Gerard Debreu, 1954, Existence of an equilibrium for a competitive economy, *Econometrica: Journal of the Econometric Society* 265–290.

Avery, Robert B, and Allen N Berger, 1991, Risk-based capital and deposit insurance reform, *Journal of Banking & Finance* 15, 847–874.

Bai, Jennie, Pierre Collin-Dufresne, R Goldstein, and Jean Helwege, 2012, Is credit event risk priced?, Technical report, Working paper.

Ben-David, Itzhak, Francesco Franzoni, Augustin Landier, and Rabih Moussawi, 2013, Do hedge funds manipulate stock prices?, *The Journal of Finance* .

Benston, George J, and George G Kaufman, 1997, Fdicia after five years, *The Journal of Economic Perspectives* 11, 139–158.

Berger, Allen N, Anil K Kashyap, Joseph M Scalise, Mark Gertler, and Benjamin M Friedman, 1995, The transformation of the us banking industry: What a long, strange trip it's been, *Brookings papers on economic activity* 1995, 55–218.

Berger, Allen N, and Gregory F Udell, 1994, Did risk-based capital allocate bank credit and cause a "credit crunch" in the united states?, *Journal of Money, credit and Banking* 26, 585–628.

Billio, Monica, Mila Getmansky, Andrew W. Lo, and Loriana Pelizzon, 2010, Econometric measures of systemic risk in the finance and insurance sectors, Working Paper 16223, National Bureau of Economic Research.

Black, Fischer, 1972, Capital market equilibrium with restricted borrowing, *The Journal of Business* 45, 444–455.

Campbell, John Y., and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* vol. 1(3), pages 195–228.

Casella, George, and Roger L. Berger, 2001, *Statistical inference (2nd ed.).* (Duxbury).

Cohen, Hugh, and William Roberds, 1993, Towards the systematic measurement of systemic risk, Working Paper 93-14, Federal Reserve Bank of Atlanta.

Davies, Sally M, and Douglas A McManus, 1991, The effects of closure policies on bank risk-taking, *Journal of Banking & Finance* 15, 917–938.

De la Fuente, Angel, 2000, *Mathematical methods and models for economists* (Cambridge University Press).

Delbaen, Freddy, and Walter Schachermayer, 1994, A general version of the fundamental theorem of asset pricing, *Mathematische annalen* 300, 463–520.

Diamond, Peter A, 1967, The role of a stock market in a general equilibrium model with technological uncertainty, *The American Economic Review* 57, 759–776.

Duffie, D., A. Eckner, G. Horel, and L Saita, 2009, Frailty correlated default, *The Journal of Finance* 64, 2089–2123.

Duffie, Darrell, and Bruno Strulovici, 2012, Capital mobility and asset pricing, *Econometrica* 80, 2469–2509.

Engle, Robert F, and Simone Manganelli, 2004, Caviar: Conditional autoregressive value at risk by regression quantiles, *Journal of Business & Economic Statistics* 22, 367–381.

Fama, Eugene F, and Kenneth R French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of financial economics* 33, 3–56.

Gai, Prasanna, and Sujit Kapadia, 2010, Contagion in financial networks, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* 466, 2401–2423.

Gale, Douglas M, and Shachar Kariv, 2007, Financial networks, *The American economic review* 97, 99–103.

Gray, D., and A. Jobst, 2010, Systemic cca - a model approach to systemic risk, Working paper, International Monetary Fund at Deutsche Bundesbank.

Gray, Dale F., Robert C. Merton, and Zvi Bodie, 2007, New framework for measuring and managing macrofinancial risk and financial stability, Working Paper 13607, National Bureau of Economic Research.

Grenadier, Steven R, 1996, The strategic exercise of options: Development cascades and overbuilding in real estate markets, *The Journal of Finance* 51, 1653–1679.

Harrison, J Michael, and David M Kreps, 1979, Martingales and arbitrage in multiperiod securities markets, *Journal of Economic theory* 20, 381–408.

Harrison, J Michael, and Stanley R Pliska, 1981, Martingales and stochastic integrals in the theory of continuous trading, *Stochastic processes and their applications* 11, 215–260.

Hovakimian, Armen, and Edward J Kane, 2000, Effectiveness of capital regulation at us commercial banks, 1985 to 1994, *the Journal of Finance* 55, 451–468.

Huang, Xin, Hao Zhou, and Haibin Zhu, 2009, A framework for assessing the systemic risk of major financial institutions, BIS Working Papers 281, Bank for International Settlements.

Humphrey, D.B., 1986, *Technology, and the Regulation of Financial Markets: Securities, Futures, and Banking.* (Lexington Books).

Ibragimov, Rustam, Dwight Jaffee, and Johan Walden, 2011, Diversification disasters, *Journal of financial economics* 99, 333–348.

Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *The review of economics and statistics* 47, 13–37.

Longstaff, Francis A., 2010, The subprime credit crisis and contagion in financial markets, *Journal of Financial Economics* 97, 436–450.

Parlour, Christine A, Richard Stanton, and Johan Walden, 2012, Financial flexibility, bank capital flows, and asset prices, *The Journal of Finance* 67, 1685–1722.

Rochet, Jean-Charles, and Jean Tirole, 1996, Interbank lending and systemic risk, *Journal of Money, Credit, and Banking* Vol. 28 No. 4.

Ross, Stephen A, 1976, The arbitrage theory of capital asset pricing, *Journal of economic theory* 13, 341–360.

Sharpe, William F, 1964, Capital asset prices: A theory of market equilibrium under conditions of risk*, *The journal of finance* 19, 425–442.

Treasury, U.S., 1991, Modernizing the financial system: recommendations for safer, more competitive banks., Technical report, U.S. Treasury Department, Washington, D.C.

Welch, Ivo, 1992, Sequential sales, learning, and cascades, *The Journal of finance* 47, 695–732.