Connections as Jumps: Estimating Financial Interconnectedness from Market Data∗

Willem J. van Vliet†

September 24, 2018

[I thank Lars Peter Hansen, Stefan Nagel, Pietro Veronesi, Azeem Shaikh, Stefano Giglio, Larry Schmidt, Bryan Kelly, Paymon Khorrami, and Mike Barnett, as well as participants of the Economic Dynamics and Financial Markets Working Group, the 2017 MFM Session for Young Scholars, the 2017 CITE Conference, and the Stevanovich Center seminar, for their helpful comments in the many iterations of this work. I am grateful for funding provided through the Liew Fama-Miller Fellowship, MFM Dissertation Fellowship, Stevanovich Fellowship, and the Bradley Foundation. All errors are my own.

†Ph.D. Candidate at the University of Chicago and Booth School of Business. Email: wvanvliet@uchicago.edu]

Abstract

I develop a new methodology for finding evidence of interconnectedness between financial institutions using readily available market price data. I argue that the classic endogeneity problem that arises when using contemporaneous price movements can be avoided by focusing on connections that trigger substantial spillovers upon default. I show that, under appropriate identification assumptions, regressing jump-like default risk on non-jump-like default risk recovers evidence of direct and indirect exposures. In my empirical work, I adapt existing techniques for estimating jump risk to a model in which the firm is a levered claim on a latent asset, and use equity, equity options, and credit default swap data on large US financial institutions to isolate jump-risk. Applying the methodology to the largest financial firms during the 2008 Financial Crisis, I find estimates of connections that are consistent with anecdotal evidence. My estimates further show that, during this period, default at connected firms would have lead to broad spillovers to the rest of the system, rather than isolated spillovers to a handful of firms. The methodology developed in the paper provides a new tool for monitoring the financial sector using contemporaneous market data.

Keywords: Financial Networks, Systemic Risk, Connectedness, Financial Crises

JEL Classification: G01, G29, C58
1 Introduction

The 2008 Financial Crisis has renewed both academic and regulatory interest in understanding connections between financial institutions. On the empirical side of this research, a new literature has evolved aiming to document which institutions are exposed to each other, and what the consequences of distress or default are. One approach to handling the multitude of channels through which firms may be connected is to use market prices to form a general measure of connectedness (Benoit et al., 2017). However, this type of approach has struggled with an endogeneity issue. Market prices are a mixture of information about the firm itself and its connections, making it difficult to isolate and identify linkages.

This paper proposes a new estimation methodology that uses contemporaneous market price changes to uncover evidence of causal linkages between firms. The key innovation is that I focus on linkages that take the form of large spillovers upon default. From an asset pricing perspective, ex ante these spillovers resemble jump risk: the risk of a sudden discontinuous jump down in the underlying assets of the financial institution. Using techniques to estimate jumps from options data, I can decompose default risk for each firm into jump risk and non-jump risk. With the identification assumption that idiosyncratic jump at one firm does not correlate with non-jump risk at another firm, finding that one firm’s non-jump risk predicts another firm’s jump risk contemporaneously is evidence of a direct or indirect causal linkage.

I apply this new methodology to data from around the 2008 Financial Crisis covering the largest financial institutions. The resulting linkages are broadly consistent with the narrative of the crisis, giving confidence that the approach may prove useful in future episodes. Furthermore, whenever a firm is linked to another firm, it tends to be linked to the majority of other firms, and not just a select few. This suggests that market participants do not believe that linkages are sparse, and instead either believe that firms are highly interconnected or that spillovers occur through broad contagion.

Focusing on large spillovers upon default, rather than a more general form of connection, is a logical approach in the context of financial firms. Implicit in the rhetoric used post-2008 Financial Crisis discussions of Systemically Important Financial Institutions and “too big/interconnected to fail” is that the primary concern to regulators is avoiding the negative consequences triggered by default, and not simply avoiding poor performance. Spillovers upon default are also the relevant type of linkage when considering a bailout. Bailouts typically only provide just enough funds or guarantees to keep a firm from defaulting, meaning that the benefit of a bailout is primarily avoiding the spillovers and potential subsequent defaults that would result from letting the default occur.

In this paper, I take no stance on what form the spillovers take. They can be direct, such as one bank directly holding the debt of another bank. They may also be indirect, such as through a contraction of funding from run-like behavior, or fire sale prices of common assets. All that matters is that the spillovers act as a large reduction in assets relative to liabilities that occurs on
the default of a connected institution. A cursory look at historical evidence does support that losses to creditors can be large when financial firms default. The failure of Lehman Brothers resulted in projected recovery rates of around 20-40% to unsecured creditors, with eventual realized payouts totaling about 35% of the outstanding face value (Scott, 2016, Table 4.1. and p. 24). Even the less complex financial firms involved in the savings and loan crisis of the 1980s and 1990s generated substantial costs. Estimates of resolution costs relative to the assets of failed savings and loan institutions peaked as high as 34.7% in 1987 (Barth, 1991, Table 3-12), implying an upper bound of 65.3% on the recovery rate for these firms.

This paper is organized in two parts. In the first, I lay out the novel identification strategy I use to find connections between financial institutions. I start under the assumption that data on financial institutions has already been transformed into the probabilities of two mutually exclusive types of defaults: defaults due to jumps (sudden, relatively rare reductions in asset values) versus defaults that are smoother (repeated small reductions in asset values), which I refer to as Brownian defaults. In a linear setup, as well in simulated data, I show that when regressing jump-like default probability at institution $i$ on all the other Brownian default probabilities, the coefficient on a particular institution $j$’s Brownian default probability is a combination of 1) $j$’s direct effect on $i$ through a linkage, 2) $j$’s indirect effect on $i$ through a chain of one or more other banks, and 3) $j$’s ability to transmit defaults arising elsewhere directly or indirectly to $i$. All three are evidence that $j$ is directly or indirectly linked to $i$, and therefore a non-zero coefficient is evidence of a direct or indirect linkage.

Key to this argument is the following identification assumption: bank-specific jump default risk unrelated to spillovers at one bank does not correlate with Brownian default risk at another bank (conditional on both aggregate jump risk and the bank’s own Brownian risk). This assumption rules out both spurious correlations and reverse causality. At the end of the paper, after presenting the empirical part, I return to this assumption. I argue that the controls used in the regression are unlikely to generate spurious, large connections.

In the second part of the paper, I show an implementation of the estimation strategy. I adapt the methodology of Andersen et al. (2015, 2016) to extract the jump intensities from equity, options, and credit default swap data on each of the financial institutions. I then estimate jump risk for the largest financial institutions during the 2008 Financial Crisis, and show that the behavior of the estimated linkages are consistent with anecdotal evidence during the crisis. Furthermore, those firms that are the source of connections tend to affect the majority of the remaining firms, rather than affecting just a few specific firms. This suggests that a broader contagion, rather than highly specific connections, was the primary concern to market participants at the time.

By utilizing a more structural approach than the previous literature, this paper makes substantial headway toward estimating causal connections using contemporaneous movements in prices. However, this approach does come with two drawbacks. First, it requires a large number of traded assets to exist for each firm, which reduces the sample to large, publicly traded financial institu-
tions. Second, it relies on extracting jump risk from all of these assets, rather than just using their prices directly. This adds a moderate level of complexity to the estimation strategy as a whole.

The approach presented in this paper also does not rely on model-based extrapolation. Instead, it relies on market participants evaluating the consequences of default and adjusting the prices of securities to reflect these consequences. The role of the model is to uncover how market participants’ expectations of spillovers are priced into securities, and how these expectations can be estimated. The benefit of this feature is that I am not extrapolating local correlations to make predictions on tail-events. The downside is that when financial institutions are relatively well capitalized, the portion of default risk that reflects the spillovers I am isolating becomes small, and is therefore overwhelmed by other sources of default risk. For example, I estimate connections using data for the two years leading up to the end of 2017, and I am unable to find any evidence of connections as the probability of a Brownian default is effectively zero.

1.1 Measuring Connections from Market Data

This paper fits in a line of literature that uses market outcomes to find evidence of connections between financial institutions. For brevity, I refer to these institutions as banks, although they need not actually be banks. Restricted to just two banks, this literature abstractly estimates equations of the form

\[(1) \quad \text{outcome}^{A} = f(\text{outcome}^{B}, \text{controls}) + \epsilon\]

where outcome\(^A\) is some outcome at bank A, outcome\(^B\) is some outcome at bank B, controls is a set of controls, and \(\epsilon\) is a residual. The types of outcomes on the left- and right-hand-sides of these equations need not be the same; they can be in different parts of a probability distribution or at different times. A non-trivial dependence of the function \(f(\cdot, \cdot)\) on outcome\(^B\) is interpreted as a sign of A being connected to B.

Without imposing additional structure, equations of the form (1) are little more than evidence of correlation. As an extreme example, simply regressing returns on contemporaneous returns recovers a scaled version of the correlation coefficient. To be of use for finding connections, the variables need to be chosen in such a way that a relationship is direct evidence of a causal relationship, or at least suggestive at some level of a causal relationship. The literature to date can be divided into two categories of approaches to finding evidence of connectedness. The first uses contemporaneous outcomes at the banks, and focuses on the tail components of outcomes to find suggestive evidence of connections. The second uses a broader set of outcomes and uses time lags to find evidence of Granger causality. My approach combines the strong points of the two: I use contemporaneous outcomes, but they are chosen in a way to rule out spurious correlation and reverse causality.
Contemporaneous Approaches  The literature that uses contemporaneous outcomes has, so far, been unable to generate causal estimates. Instead, these approaches find suggestive features of the joint distribution of outcomes. Though not causal, these estimates are nonetheless interesting as they highlight features of the tails of the distributions involved.

A prominent example of this approach is the ∆CoVaR measure of Adrian and Brunnermeier (2016). When restricted to just two banks, ∆CoVaR is defined to be how much the conditional value-at-risk\(^1\) (CoVaR) at bank A increases when we change from conditioning on a median return at bank B to a 95th-percentile bad return at B. As the authors note in their paper, it is important to realize that the movement in conditioning information need not be exogenous, and hence the change in the CoVaR cannot generally be seen as causal. The two banks may simply be correlated, in which case knowing that bank B had a bad return increases the likelihood that bank A had one as well. Even worse for determining a directional causal relationship, it could be bank B that is exposed to bank A, and hence bank B’s bad return could be reflective of bank A’s bad return.

Nonetheless, the ∆CoVaR measure has intuitive appeal: it tells us how much the tail of the distribution of returns at bank A moves when bank B moves from its median to its tail. It therefore focuses on a relevant part of the distribution for the affected bank, and is not mechanically symmetric. In my estimation strategy, I build upon this insight of using different parts of the distribution further by isolating portions of the tail attributable to jump risk.

∆CoVaR is not the only measure based on contemporaneous outcomes. The Co-Risk measure of International Monetary Fund (2009), based on an earlier version of Adrian and Brunnermeier (2016), performs a similar exercise in the context of credit default swap spreads. Goodhart and Segoviano Basurto (2009) estimates the joint distribution of bank outcomes, and then computes the distribution dependence matrix (probability of distress at one bank conditional on distress at another particular bank) and the probability of cascade effects (probability of one or more other banks being in distress conditional on a particular bank being in distress). Giglio (2014) uses counterparty risk in credit default swaps to compute bounds on the joint probability of default, and then computes the probability that a particular bank is involved in a simultaneous multiple bank default. As with ∆CoVaR, these papers all provide interesting insights about the responses of tail end variables, but none of them are able to, or claim to, provide causal estimates.

Causal Approaches  The literature aimed at uncovering causal relationships between financial institutions relies on Granger causality to achieve identification. These papers label bank A as being connected to bank B if outcomes at B Granger cause outcomes at A. An informative example in this strand is the Granger causality measures of Billio et al. (2012). After some scaling, and

\(^1\)The \(q\)th percentile value-at-risk (VaR) is defined to be the \(q\)th percentile of the bank’s return loss distribution. \(q\) is typically chosen to be 95%. The \(q\)th percentile conditional value-at-risk (CoVaR) is the \(q\)th percentile of the bank’s return loss distribution conditional on some specified event.
potentially controlling for aggregate factors, the test is effectively to run the regression

\[ r_{t+1}^{[A]} = \alpha + \beta r_t^{[A]} + \gamma r_t^{[B]} + \epsilon_{t+1} \]

where \( r_t^{[x]} \) denotes the equity return at bank \( x \) at time \( t \). Bank A is connected to bank B if bank B’s returns Granger-cause bank A’s returns (\( \gamma \neq 0 \)). With Granger causality, the usual intuition is that information cannot flow backwards in time, and hence \( \gamma \neq 0 \) is not a reflection of the future return at bank A affecting the current return at bank B.

In the context of financial data, however, we may not expect to find much Granger causality even in the presence of strong connections, and non-zero coefficients may not reflect a readily-interpretable type of influence. In a model with rational expectations and with agents perfectly informed about any connections, any information about bank B at time \( t \) should be priced into bank A contemporaneously. When looking at returns data in such a setup, the effect of \( r_t^{[B]} \) should be contained entirely in \( r_t^{[A]} \), with \( r_{t+1}^{[A]} \) remaining independent. Appealing to Granger causality therefore implicitly requires a departure from rational expectations, a slow dispersion of information about connections to agents in the economy, or some other type of friction to delay the incorporation of information.

Other examples in this literature are as follows. Basu et al. (2017) directly extends Billio et al. (2012) by estimating a full vector autoregression of returns instead of performing pairwise regressions, and uses “debiasing” Lasso techniques coupled with controlling the false discovery rate to perform inference in a high-dimensional setting. Diebold and Yılmaz (2014, 2015) consider equity return volatilities instead of returns, and use forecast error variance decompositions to find evidence of linkages in a joint estimation. Demirer et al. (2017) extend this work by looking at global banks, and also uses Lasso to handle the large number of regressors. Like Billio et al. (2012), these papers use information that is revealed over time to document connectedness in financial markets, meaning that these approaches face the same limitations.

To obtain causal estimates, my estimation strategy also relies on finding outcomes with a unidirectional information flow. I detect connections by finding cases where non-jump-like default probability at one bank predicts jump-like default probability at another. As will be discussed at length in Section 3, the key identification assumption is that idiosyncratic jump-like default risk does not cause or correlate spuriously with non-jump-like default risk at other banks. This parallels the identification assumption used in Granger causality, where one assumes that information cannot

---

2Technically, their measure is a mixture of Granger causality and covariance. The Diebold-Yılmaz approach relies on generalized variance decompositions, which includes both the dynamic effects and the contemporaneous effects as given by the covariance matrix of shocks. To avoid making shock identification assumptions, the generalized variance decomposition works by considering generalized impulse responses. When shocking a particular variable in the impulse response, all other variables are shocked by their conditional expected value. In the normal case considered, this is just the best linear predictor. As a result, any correlation between shocks will tend to lead to both variables playing a role in each other’s generalized variance decomposition. For the more aggregated measure of connectedness, the authors do consider different shock identification strategies. However, as these are different Cholesky orderings, they effectively do the exact opposite of the generalized variance decomposition. This approach posits that any (contemporaneous) connection must be unilateral.
flow backward in time, i.e., future outcomes cannot cause current outcomes. Unlike with time, there is no obvious barrier preventing the backwards flow of information, and therefore I analyze the robustness of the identification assumption in depth in Section 6.

1.2 Other Related Literature

The literature on estimating the interconnectedness of the financial system is wider than the examples already discussed. Broadly speaking, this literature can be divided into two approaches (Benoit et al., 2017). The first, which this paper and the previous examples all belong to, uses market data to find evidence of connections or systemic risk. As already discussed, these papers allow for broad notions of connectedness, but must deal with the difficulty of endogeneity. In addition to Benoit et al. (2017), an overview of these approaches can be found in Bisias et al. (2012).

The second approach selects a particular channel of connections, and often uses proprietary or confidential data to estimate or document this channel. The benefit of this approach is that for each particular mechanism it is often possible to estimate directional and causal linkages. Furthermore, by specifying a particular channel, this approach enables precise counterfactual exercises. The drawback is that estimates are confined to a particular channel, and therefore may only present a portion of the overall level of connectedness. Hüser (2015) provides an additional survey that focuses primarily on this type of approach.

This paper addresses a fundamentally different question than some popular measures of systemic risk, and should not be confused with these approaches. Many advances in measuring systemic risk focus on identifying which institutions contribute to systemic risk. Popular examples of this type are SRISK (Brownlees and Engle, 2017), marginal/systemic expected shortfall (MES/SES, Acharya et al., 2017), and the distress insurance premium (DIP, Huang et al., 2012). These papers concern computing shortfalls or losses at each individual firm conditional on an aggregate shortfall or loss, and can be thought of as a decomposition of an aggregate event into its constituents. My paper addresses a different question; instead of decomposing systemic shortfalls into its components, I estimate whether defaults during systemic times are likely to lead to further defaults.

To estimate jump risk, this paper combines the work of two separate literatures. As I use both equity- and debt-related asset prices, I use a Merton (1974) style model where the firm is treated as a levered latent asset. Price are then computed by treating the market assets as options on the latent asset.\(^3\) To estimate jump risk for the underlying model, I adapt the implementation of Andersen et al. (2016) to work with a latent asset.

---

\(^3\) Merton (1974) is the first work to apply this approach, although Black and Scholes (1973); Merton (1970, 1973) all point toward this approach.
2 A Conceptual Framework of Defaults and Connectedness

Before discussing the empirical identification strategy, it is helpful to understand when banks default if their assets are exposed to diffusive shocks, jumps, and spillovers from other banks. To that end, I first develop a formal model where I can rigorously define four types of default risk: Brownian risk, idiosyncratic jump risk, aggregate jump risk, and spillover risk. Each of these types of risk correspond what caused the bank to default: a sufficiently negative diffusive path, a negative idiosyncratic jump, a negative aggregate jump, or a spillover from another bank. Once defined, I linearly approximate the model to better understand how spillover risk depends on default risks at the other banks. This relationship is then used extensively when discussing identification.

2.1 Formal Model

To be more concrete, suppose we have $N$ banks, and time is continuous. In the style of Merton (1974), I treat each bank as a levered asset, and denote the value of assets relative to liabilities at bank $i$ at time $t$ by $A_t[i]$. Bank $i$ defaults when $A_t[i]$ reaches or crosses 1 from above.

Assets relative to liabilities at banks follow a jump-diffusion process when they are all away from their respective default boundaries. If $A_t[i] > 1$ for all $i$, then each bank evolves according to

$$\frac{dA_t[i]}{A_t[i]} = \beta_B[i] \sigma_t dZ_t + \beta_J[i] \left( d \left( \sum_{j=1}^{N_t} (V_j - 1) \right) + \mathbb{E}[1 - V] \lambda_t dt \right)$$

$$+ \sigma_t[i] dZ_t[i] + d \left( \sum_{j=1}^{N_t[i]} (V_{ij} - 1) \right) + \mathbb{E}[1 - V[i]] \lambda_t[i] dt$$

(2)

Here, $Z_t$ is a standard Brownian motion that captures aggregate Brownian (diffusive) shocks, and is common to all banks. $\sigma_t$ is the instantaneous volatility of the aggregate Brownian process, and $\beta_B[i]$ is bank $i$’s exposure to the aggregate Brownian shock. Similarly, $Z_t[i]$ is an independent standard Brownian motion that captures idiosyncratic Brownian shocks, with $\sigma_t[i]$ denoting the instantaneous volatility. $N_t$ is an independent, inhomogeneous Poisson process with instantaneous arrival rate $\lambda_t$ that counts the number of aggregate shocks that have occurred. Each $V_j$ is an i.i.d. aggregate jump size. The term $\mathbb{E}[1 - V] \lambda_t$ is a jump compensator. $\beta_J[i]$ is bank $i$’s exposure to the aggregate jump process. $N_t[i]$ is an independent, inhomogeneous Poisson process with instantaneous arrival rate $\lambda_t[i]$ that counts the number of bank-specific jumps that have occurred. These jumps have i.i.d. jump size $V_{ij}$, and the term $\mathbb{E}[1 - V[i]] \lambda_t[i]$ is a jump compensator. I do not make any independence assumptions on the idiosyncratic pieces across banks.\(^4\)

Spillovers from default are modeled as a boundary condition to the system as a whole. When one or more banks hit or cross their default boundaries, their defaults may spill over to other banks.

\(^4\)For example, $dZ_t[i]$ can be correlated with $dZ_t[j]$ for $i \neq j$ and the jumps can be interdependent as well.
and cause subsequent defaults. Let $\theta_{i,j} \geq 0$ denote the reduction in assets that bank $i$ incurs when bank $j$ defaults. For each bank $i$, the value of assets $A_{i+}^{[i]}$ after accounting for all the spillovers solves the equation

$$A_{i+}^{[i]} = A_{i}^{[i]} - \sum_{j \neq i} \theta_{i,j} 1 \{ A_{i+}^{[j]} \leq 1 \}$$

Given the values $A_{i+}^{[j]}$ for all $j \neq i$, equation (3) gives a unique value for $A_{i+}^{[i]}$. However, when stacked across all banks, equation (3) becomes a fixed-point problem that potentially has multiple solutions. Multiplicity arises from what are essentially “self-fulfilling defaults”: a group of banks start off with assets above liabilities, but they all default if the others default. For simplicity, I choose the unique solution to equation (3) that rules out these self-fulfilling defaults. I discuss this solution in more detail in Appendix B. I also discuss the effect of allowing self-fulfilling defaults in Appendix B.1, and argue that a different choice of equilibrium does not substantially alter the results.

Once the $A_{i+}^{[i]}$ values are computed, banks that have defaulted ($A_{i+}^{[i]} \leq 1$) are removed. They no longer cause any subsequent spillovers. The remaining banks, which have assets above liabilities, continue to evolve according to their respective jump-diffusion processes.

From bank $i$’s perspective, its assets effectively follow a jump-diffusion with three sources of jumps. The first two are the aggregate and idiosyncratic jumps already in equation (2). The third is spillovers from other banks defaulting. Though spillovers are an endogenous outcome of the system, and not a random process to the system as a whole, from bank $i$’s perspective these spillovers are still probabilistic discontinuous reductions in asset value.

### 2.2 Decomposition of Default Risk

The formal model presented above has four pathways to default, corresponding to the four sources of risk to the assets of each bank. Given a default time horizon, which I denote $\tau^D > 0$, this leads to four mutually exclusive default events. These four events allow us to decompose the total probability of default into the four corresponding parts.

The first, which I will refer to as Brownian risk, is the probability of defaulting due to a sufficiently negative diffusive path. Brownian risk captures the case where at the time of default the assets did not experience any type of jump, and assets hit the default boundary prior to any spillovers. Note that Brownian risk includes both the contribution from the idiosyncratic diffusive shocks ($dZ_{i+}^{[i]}$) and aggregate diffusive shocks ($dZ_t$) in equation (2), as these are indistinguishable.

The next two types of defaults are from the two sources of jump risk in equation (2). Idiosyncratic jump risk is the probability that an idiosyncratic jump causes the default at the bank. Likewise, aggregate jump risk is the probability that an aggregate jump causes the default at the bank. Both types of defaults are characterized by assets being strictly above the default boundary,
and then experiencing a discontinuous jump down to or below the boundary prior to any spillovers.

Finally, there is spillover risk, which is the probability of defaulting as a result of a spillover. In this case, the bank’s assets remain above the default boundary prior to computing any spillovers, but after any spillovers given by equation (3) have been realized, the bank’s assets are below the default boundary.

Table 1 formalizes these categories in the mathematical notation of the model, and introduces the names I will use for them in formulas below. Note that as the events for each of the probabilities are mutually exclusive, the default probabilities simply add up to form probabilities of unions of events. In particular, the sum of all four is the total default probability in the next $\tau^D$ units of time.

### 2.3 Linear Approximation

Even for very simple stochastic processes, computing the default probabilities of Section 2.2 in the formal model of Section 2.1 is intractable. Spillovers are endogenous jumps in the system, and are discontinuous boundary conditions rather than additional terms in the differential equations characterizing the relevant value functions. As a result, it is difficult to characterize the dependence of default probabilities directly. Instead, I work with a linearized version of the model.

In Appendix C, I linearize an approximation to the formal model in the two bank case. I show that, apart from some corner cases, the model behaves as would be expected: if bank 1 is connected to bank 2, then an increase in bank 2’s idiosyncratic jump default probability or Brownian default probability results in an increase in the spillover default probability at bank 1. When bank 1 is not connected to bank 2, the dependence vanishes. Furthermore, bank 1’s spillover default probability can depend positively or negatively on its own Brownian and idiosyncratic default probabilities. The dependence tends to be positive when increases in these default probabilities also signal an increase in the likelihood that a spillover leads to a default, which I refer to as sensitivity. The dependence tends to be negative when these own defaults precede any spillover, which I refer to as a race condition.

To analyze how the network generates spillover risk for the full $N$ bank case, I start with a linear setup guided by the two bank case. For simplicity, I put a common coefficient on Brownian default risk, idiosyncratic default risk, and aggregate jump risk.\footnote{This assumption for aggregate jump risk is inconsequential, as future regressions control for aggregate jumps. The assumption that the coefficient on idiosyncratic jump risk is the same as the coefficient on Brownian results in the particularly simple expressions for Assumption 2 and Proposition 1. With different coefficients, the lower bound on the covariance changes, but the intuition remains unchanged.} To a first order approximation,
this yields

\[
\text{spillover}_t^{[i]} \approx \sum_{j \neq i} c_{j}^{[i]} \text{non\_spillover}^{[j]}_t + \sum_{j \neq i} d_{j}^{[i]} \text{spillover}^{[j]}_t + f^{[i]} \text{non\_spillover}^{[i]}_t
\]

where \(\text{non\_spillover}^{[j]}_t \equiv \text{Brownian}^{[j]}_t + \text{agg\_jump}^{[j]}_t + \text{idio\_jump}^{[j]}_t\) is simply the probability of defaulting in the next \(\tau_D\) units of time due to any reason except for a spillover from another bank.

The coefficients \(c_{j}^{[i]}\) and \(d_{j}^{[i]}\) capture the threat of a future default at bank \(j\) spilling over to bank \(i\) through a direct link (\(\theta_{i,j} > 0\)). Both coefficients are nonnegative. The coefficient \(c_{j}^{[i]}\) therefore captures how defaults that originate at bank \(j\) affect bank \(i\), while \(d_{j}^{[i]}\) captures how defaults that originate elsewhere but have led to a default at \(j\) subsequently affect \(i\). These two effects need not be the same, and for this reason I have parameterized them separately. For example, if bank 3 has spillovers to both banks 1 and 2, and bank 2 has spillovers to bank 1, the coefficient \(d_{2}^{[1]}\) only captures the marginal effect of bank 2 on bank 1 when bank 3 defaults. \(d_{2}^{[1]}\) is only captures the cases where bank 1 survives the direct spillover from bank 3, but then defaults because of the additional spillover from bank 2. By contrast, \(c_{2}^{[1]}\) captures the entire of bank 2 on bank 1 when bank 2 defaults on its own.

As a result of this distinction, \(c_{j}^{[i]}\) is nonzero if and only if \(\theta_{i,j} > 0\), while with \(d_{j}^{[i]}\) we only have that if \(d_{j}^{[i]}\) is nonzero then \(\theta_{i,j} > 0\). Either way, a nonzero value of \(c_{j}^{[i]}\) or \(d_{j}^{[i]}\) is evidence of \(\theta_{i,j} > 0\).

The coefficient \(f^{[i]}\) in equation (4) captures two different ideas. One the one hand, the more likely bank \(i\) is to default on its own, the more likely its default will precede any spillovers from other banks. This race between defaulting on its own and receiving a spillover tends to push \(f^{[i]}\) to be negative. On the other hand, the more likely bank \(i\) is to default on its own, the more likely it is to be closer to its default boundary, resulting in the bank being more sensitive to spillovers from other banks. This sensitivity tends to push \(f^{[i]}\) to be positive. Whichever force dominates determines the sign of the coefficient.

Stacking equation 4) across all the banks allows us to solve for the spillover default probabilities as a linear function of the non-spillover default probabilities. Let bold probabilities denote stacked versions of the probabilities. Let bold, capital coefficients denote the matrix versions of coefficients, and bold, lowercase coefficients denote the diagonal matrix versions of coefficients. For example, \(\text{spillover}_t = (\text{spillover}_t^{[1]} \cdots \text{spillover}_t^{[N]})', \ C = (c_{j}^{[i]}), \ f = \text{diag}(f^{[1]} \cdots f^{[N]}))\). Then

\[
\text{spillover}_t = (I - D)^{-1}(C + f) \text{non\_spillover}_t
\]

Equation (5) gives us the spillover default probabilities in terms of non-spillover default probabilities, and will play a central role in identification. The simplicity relied on the assumption that all types of non-spillover default risk spill over equally to other banks. If we relax this assum-
tion, equation (5) would instead feature three such terms, one for all Brownian risk, one for all idiosyncratic jump risk, and one for aggregate jump risk. The intuition would be the same, and the identification assumptions would be unaffected.

Interpreting Equation (5)  The $i$th column of the coefficient $(\mathbb{I} - D)^{-1}(C + f)$ in equation (5) captures the cumulative effect of raising the non-spillover default probability at bank $i$ on all other banks. It can be decomposed into four parts

$$ (\mathbb{I} - D)^{-1}(C + f) = C + ((\mathbb{I} - D)^{-1} - \mathbb{I}) C + f + ((\mathbb{I} - D)^{-1} - \mathbb{I}) f $$

The first term is the effect coming through direct exposures, and is simply the first term in equation (4) in matrix form. The second term, which can also be written as $DC + D^2C + D^3C + \cdots$, captures indirect exposures: $DC$ captures how defaults from spillovers arising from non-spillovers propagate through the network, $D^2C$ captures how these subsequent defaults propagate, etc. The third term is the changing susceptibility of each bank to spillovers, whether through defaulting on its own before a spillover can happen or becoming more likely to default if a spillover happens. The fourth term, which can be written as $Df + D^2f + \cdots$, is how this changing susceptibility to spillovers affects banks further down the line. The fourth term therefore can be thought of as capturing the changing role of each bank in transmitting shocks to other banks.

Column $i$ of the total coefficient $(\mathbb{I} - D)^{-1}(C + f)$ therefore captures a mixture of 1) the direct transmission of defaults that originate at bank $i$, 2) the indirect transmissions of defaults originating at bank $i$, and 3) the changing role of bank $i$ in transmitting shocks to each bank. The $(i, i)$ element additionally captures bank $i$’s susceptibility to defaulting from a spillover. Importantly, for $j \neq i$, a non-zero $(j, i)$ entry in the coefficient indicates that bank $j$ is directly or indirectly connected to bank $i$.

Self-Amplification  Note that equation (5) contains a mistake from linearization. The quantity $(\mathbb{I} - D)^{-1}C$ may have a non-zero diagonal, which taken literally would mean that some bank $i$’s default probability rises more than one-for-one as a result of an increase of its own default probability. Obviously, unless we are considering self-fulfilling default equilibria, such amplification is impossible: the network does not make a default at bank $i$ in one state of the world cause a default at bank $i$ in another state of the world. In Appendix D I consider a more complicated solution that does not yield the same self-amplification. By explicitly considering all paths that defaults can take through the network, I avoid the non-zero diagonal. While the solution is no longer as simple, the interpretation of the coefficient remains unchanged.
3 Identification

The conceptual decomposition of default probabilities in the formal model of Section 2 are useful for understanding the effects of spillovers, but the decomposition is finer than can be achieved using data. The two sources of jump risk and the spillovers all appear as discontinuities in the value of assets relative to liabilities, and can therefore not be separated using price data alone. From data we can only obtain a combination of these three jump-like default probabilities, as well as the Brownian default probability and a proxy for an aggregate jump probability. In this section, I show that, under appropriate identification assumptions, the cumulative coefficient that captures both direct and indirect effects can be estimated using these values. Evidence of linkages are therefore identified, despite not being able to isolate the spillover default probability.

For notation, Brownian is the stacked vector of Brownian default probabilities. \( \text{jump}_i \) is the probability of any of the three jump-like events occurring at bank \( i \), with \( \text{jump}_i \) being the stacked vector across all banks. \( \hat{\text{agg}} \text{jump}_t \) is a scalar estimate of the aggregate jump probability. These quantities are the only ones that can be extracted from data.

3.1 Regression

Breaking apart \( \text{non_spillover}_i \) in equation (5) into its three constituent parts, and then adding \( \text{agg}_i \text{jump}_i + \text{idio}_i \text{jump}_i \) to both sides yields

\[
\text{jump}_i = (\mathbb{I} - D)^{-1}(C + f) \cdot \text{Brownian}_t
\]

\[
+ (\mathbb{I} + (\mathbb{I} - D)^{-1}(C + f)) \cdot \text{agg}_i \text{jump}_i
\]

\[
+ (\mathbb{I} + (\mathbb{I} - D)^{-1}(C + f)) \cdot \text{idio}_i \text{jump}_i
\]

The first term consists only of observable quantities. The second term contains the unobserved aggregate jump probabilities at each bank, but each of these can be proxied for by using \( \hat{\text{agg}} \text{jump}_t \). The third term is unobserved with no proxy.

Equation (6) suggests that the coefficient \( (\mathbb{I} - D)^{-1}(C + f) \) can be estimated row-by-row by regressing each of the jump-like default probabilities on all of the Brownian default probabilities and the proxy for aggregate jump risk. The remaining idiosyncratic jump risk would then be subsumed in the residual for such a regression. To allow for non-stationary drivers of the variables, or to approximate an autoregressive environment with high levels of autocorrelations, I therefore consider this regression in first differences:

\[
\Delta \text{jump}_i = \alpha_i + \beta_i \Delta \text{agg}_i \text{jump}_t + \chi_i \Delta \text{Brownian}_t + \sum_{j \neq i} \gamma_{i,j} \Delta \text{Brownian}_j + \epsilon_{i,t+1}
\]

Recall that \( (\mathbb{I} - D)^{-1}(C + f)_{i,j} \) captures the direct and indirect effects of bank \( j \) on bank \( i \),

12
and therefore a non-zero value is evidence of a direct or indirect linkage from bank \( j \) to \( i \). If the population coefficient \( \gamma_{ij}^\ast \) is proportional to \((I - D)^{-1}(C + f)_{i,j}\), then a non-zero \( \gamma_{ij}^\ast \) is evidence of a direct or indirect linkage from bank \( j \) to bank \( i \). I now turn to the assumptions needed on the unobserved idiosyncratic jump risk needed to prevent omitted variable bias from generating spurious results.

### 3.2 Identification Assumption

The usual condition to prevent omitted variable bias is that the omitted variables are orthogonal to the regressors. In regression (7) this is simply given by the following assumption:

**Assumption 1.** The (unobserved) idiosyncratic jump intensity does not correlate with Brownian risk, i.e.

\[
\text{cov}(\Delta \text{idio\_jump}_{i+1}^j, \Delta \text{Brownian}_{i+1}^i) = 0
\]

for all \( i, j \).

Under this assumption, and the usual regularity conditions for regression, the regression coefficients \( \gamma_{ij}^\ast \) recovers the \((i,j)\) element of \((I - D)^{-1}(C + f)\). As discussed in Section 2.3, this is a combination of the direct and indirect effects of spillover risk from bank \( j \) to bank \( i \), as well as bank \( j \)’s role in transmitting spillovers originating elsewhere.

Assumption 1 also must hold for \( i = j \), i.e., that Brownian and idiosyncratic jump risk within a particular bank are uncorrelated. This condition may be too strong; a bank that is behaving worse may simultaneously have more Brownian and more jump risk. However, we can relax this requirement and still leave the sign of the estimated coefficient unchanged under the following assumption.

**Assumption 2.** The following two conditions hold:

(i) The (unobserved) jump intensity does not correlate with Brownian risk across banks beyond dependence on its own Brownian risk

\[
\text{cov}\left(\Delta \text{idio\_jump}_{t+1}^j, \Delta \text{Brownian}_{t+1}^i \mid \Delta \text{Brownian}_{t+1}^j, \Delta \text{agg\_jump}_{t+1}\right) = 0 \quad \text{for all } i \neq j
\]

(ii) Within each bank, the covariance of Brownian risk and idiosyncratic jump risk, conditional on aggregate jump risk, is not too negative

\[
\frac{\text{cov}\left(\Delta \text{idio\_jump}_{t+1}^i, \Delta \text{Brownian}_{t+1}^i \mid \Delta \text{agg\_jump}_{t+1}\right)}{\text{var}\left(\Delta \text{Brownian}_{t+1}^i \mid \Delta \text{agg\_jump}_{t+1}\right)} > -1 \quad \text{for all } i
\]

In both conditions, \( \text{cov}(x, y \mid z) \) indicates the covariance between \( x \) and \( y \) after conditioning linearly on \( z \).
This weakened assumption effectively allows for some correlation between the idiosyncratic jump probabilities and the Brownian probabilities, but only through each bank’s own Brownian probability. Equivalently, it allows a linear relationship $\Delta \text{idio}^{\text{jump}}_{i,t+1} = \rho^{[i]} \Delta \text{Brownian}^{[i]}_{i,t+1} + \omega^{[i]} \Delta \text{agg}^{\text{jump}}_{t+1} + v^{[i]}_{t+1}$ where $v^{[i]}_{t+1}$ is orthogonal to all the Brownian variables and the aggregate jump probability, with $\rho^{[i]} > -1$. In this case, we obtain scaled versions of the coefficients obtained under Assumption 1.

**Proposition 1.** Under Assumption 2, and the usual regularity conditions for regression, the population coefficient

$$
\gamma^{[i]}_{j} = \left(1 + \frac{\text{cov} \left( \Delta \text{idio}^{\text{jump}}_{j,t+1}, \Delta \text{Brownian}^{[j]}_{j,t+1} \mid \Delta \text{agg}^{\text{jump}}_{t+1} \right)}{\text{var} \left( \Delta \text{Brownian}^{[j]}_{j,t+1} \mid \Delta \text{agg}^{\text{jump}}_{t+1} \right)} \right) \left( (\mathbb{I} - D)^{-1} (C + f) \right)_{i,j}
$$

In particular, it has the same sign as $( (\mathbb{I} - D)^{-1} (C + f) )_{i,j}$.

**Proof.** See Appendix E.2.

The intuition for this result is straightforward, and results from idiosyncratic jump risk and Brownian risk having the same consequences. The estimated coefficient then picks up a blend of the two, but as they are both evidence of the same linkage, the blend still is evidence of the same linkage. As long as the idiosyncratic jump piece does not offset the Brownian piece (the second condition in Assumption 2), the coefficient is still positive and still reflects a direct or indirect linkage. In fact, the formula for the coefficient in Proposition 1 is still correct even if the covariance between the idiosyncratic jump risk and Brownian risk at $j$ is too negative. In this case, the sign changes, but it still remains true that a non-zero coefficient is evidence of a linkage.

One of Assumptions 1 and 2 is required for non-zero coefficients to be evidence of causal linkages. Either assumption implies that a non-zero coefficient $\gamma^{[i]}_{j}$ is evidence of a direct or indirect linkage from bank $j$ to bank $i$. The stronger Assumption 1 further gives us that the magnitude of $\gamma^{[i]}_{j}$ is the relative increase in the spillover default probability at bank $i$ in response to an increase in the non-spillover default probability at bank $j$.

Without either of the assumptions holding, nonzero coefficients may simply be uncovering a correlation between the Brownian risk of one bank and the idiosyncratic risk of another bank. Even more worryingly, nonzero coefficients could be evidence of reverse causality, where it is jump risk from one bank that causes Brownian risk at the other. I discuss violations of these two assumptions in more detail in Section 6. In the broad set of cases I consider, I argue that violations are unlikely to generate large, spurious coefficient estimates.
3.3 Simulation Results

With all the moving parts, it can be helpful to see a numerical illustration of the identification strategy. To that end, I now show the results of running the various regressions discussed in simulated data. For computational simplicity, I use $N = 3$ banks, as this is the minimal number of banks needed to show each of the channels.

For simplicity, I fix $\sigma_t$ and $\sigma^i_t$ to be constant. I specify the aggregate jump size to be log-normal $V_j \sim \log N(\mu_J, \sigma_J^2)$ and the idiosyncratic jump size to also be log-normal $V^i_j \sim \log N(\mu^i_J, \sigma^i_J^2)$. The arrival rates for the jump processes are modeled as Cox, Ingersoll and Ross (1985) (CIR) processes (more broadly known as a Feller (1951) square-root process). The aggregate jump intensity is given by
\[
d\lambda_t = \kappa_\lambda (\bar{\lambda} - \lambda_t)dt + \sigma_\lambda \sqrt{\lambda_t}dW_t
\]
where $W_t$ is an independent standard Brownian motion, and the idiosyncratic jump intensity for bank $i$ is
\[
d\lambda^i_t = \kappa^i_\lambda (\bar{\lambda}^i - \lambda^i_t)dt + \sigma^i_\lambda \sqrt{\lambda^i_t}dW^i_t
\]
where $W^i_t$ is an independent standard Brownian motion.

Table 2 shows an example for moderate parameters. All banks in isolation are identical, and only the spillovers differ between each bank. Defaults at bank 3 spill over to bank 2, and defaults at bank 2 spill over to bank 1. Note that this means that bank 1 is indirectly connected to bank 3 through bank 2. Panel (a) documents all of the parameters and initial values chosen.

Panel (b) show the results of regression (7). The lower triangle is significant, while the upper triangle is near zero and insignificant. Recall that coefficients represent either direct or indirect linkages, with no way to distinguish between the two cases. The coefficients from bank 3 to bank 2 and from bank 2 to bank 1 reflect the direct linkages, while the positive coefficient from bank 3 to bank 1 is as a result of the indirect linkage through bank 2. The negative coefficients on bank 1’s and bank 2’s own Brownian risk highlight that here the race condition mentioned earlier for $f^i$ dominates. This example has a relatively large common Brownian component to default risk, and therefore periods where spillovers happen are also periods where the bank is likely to default from a bad Brownian shock. The negative coefficients captures the reduction in the chance of a spillover happening. By contrast, bank 3 is not exposed to any other banks, and hence there is less of a horse race between jump risk and Brownian risk.

I show several different parameterizations in Appendix F. The coefficients behave as would be expected; when Brownian volatilities increase or the size of spillovers decrease, the size of the coefficients decrease. The simulations performed do highlight an interesting feature of the model that is difficult to show analytically but nonetheless intuitive. Even with relatively small spillovers,
the coefficients can still be sizable if the correlation among banks is high. The reason for this is that with high correlation all banks tend to be closer to their default boundaries at the same time. When one bank fails, the others do not have large buffers, and even a small spillover can drive them to default. When correlations are lower, the other two banks tend to have a larger buffer, and therefore tend to survive the small spillovers, resulting in smaller coefficients.

4 Extracting Jump Risk

The identification arguments presented so far rely on jump-like and Brownian default probabilities extracted from data, as well as a measure of aggregate jump risk. In this section, I present a method of extracting these components from observable credit default swap (CDS) spreads, options prices, and equity prices. I start with a Merton (1974) style model of each bank, and price all of these assets within the context of the model. I then present the modification of Andersen et al. (2016) that I use, as well as the modification of the volatility estimate used. I also discuss the modifications for extracting aggregate jump risk from S&P500 data, and I discuss the data sources used.

4.1 Asset Prices

As they are statistically indistinguishable at any single bank, I collapse the aggregate and idiosyncratic components of equation (2) into one. The two types of Brownian and jump shocks are useful for exposition, but are difficult to distinguish in the data. The latent asset \( A_i^{[i]} \) of bank \( i \) now has dynamics under \( \mathbb{Q} \)

\[
\frac{dA_i^{[i]}}{A_i^{[i]}} = r_t dt + \sigma_i^{[i]} dZ_i^{[i],Q} + d \left( \sum_{i=1}^{N_i^{[i]}(t)} (V_i - 1) \right) + \lambda_i^{[i]} \xi_i^{[i]} dt
\]

where \( r_t \) is the risk-free rate, \( \sigma_i^{[i]} \) is the instantaneous volatility, \( N_i^{[i]}(t) \) is an inhomogeneous Poisson process with intensity \( \lambda_i^{[i]} \), \( V_i \sim \log N(\mu_i^{[i]}, \sigma_i^{[i]}^2) \) is a log-normal jump size, and \( \xi_i^{[i]} = 1 - \exp\{\mu_i^{[i]} + \frac{1}{2} \sigma_i^{[i]}^2\} \) is a jump compensator. Here, I assume \( \mu_i < 0 \) and \( \xi_i^{[i]} > 0 \); the standard deviation of the log-normal distribution is not so large that jumps become positive in expectation.

Rather than specifying the dynamics for \( r_t, \sigma_i^{[i]}, \text{ and } \lambda_i^{[i]} \), I “freeze” each of these state variables to form a short-term approximation as in Andersen et al. (2016). When looking at assets priced at date \( t \), I let \( A_{s \geq t}^{[i]} \) evolve as if \( r_s, \sigma_s^{[i]}, \text{ and } \lambda_s^{[i]} \) are all fixed. The approximation to (8) is

\[
\frac{dA_s^{[i]}}{A_s^{[i]}} = r_t ds + \sigma_i^{[i]} dZ_s^{[i],Q} + d \left( \sum_{i=1}^{N_i^{[i]}(s)} (V_i - 1) \right) + \lambda_i^{[i]} \xi_i^{[i]} dt
\]

Now, \( N_i^{[i]}(s) \) is a homogeneous Poisson process with fixed intensity \( \lambda_i^{[i]} \).

16
The benefit of equation (9) is that the distribution of $A_s | A_t$ has a closed form for all $t \geq s$. It is an (infinite) mixture of log-normal distributions, each corresponding to the number of jumps that have occurred between time $t$ and $s$. The mixture weights come from the Poisson distribution with rate parameter $\lambda^i_t (s - t)$.

I treat assets on the bank $i$ as options on $A^i_t$, in the style of Merton (1974). Given a default time horizon $\tau_D$, I treat the firm at time $t$ as a levered debt claim on $A^i_{t+\tau_D}$, where debt has a face value of 1. The value of equity is the residual value, and hence at time $t$ the value of equity relative to liabilities is the value of a call option on $A^i_{t+\tau_D}$ with strike price 1.

In general, an equity option with maturity less than $\tau_D$ will be a compound option. However, by considering options of the same maturity as debt ($\tau_D$), the compound options collapse to options and option spreads. Specifically, a call option at time $t$ on equity relative to liabilities with strike $k$ and maturity $\tau_D$ becomes a call option on $A^i_{t+\tau_D}$ with strike $1 + k$. A put option at time $t$ on equity relative to liabilities with strike $k$ and maturity $\tau_D$ becomes a put option spread on $A^i_{t+\tau_D}$ that is long a put option with strike price $1 + k$ and short a put option with strike price 1.

Lastly, I will be using credit default swap (CDS) data. A credit default swap can be thought of as a stream of payments (the spread) from the protection buyer to the protection seller that continue until either the contract expires or the reference security experiences a credit event. In the context of the model, this would imply that the CDS spread at time $t$ is intimately tied to the value of a put option on $A^i_{t+\tau_D}$ with strike price 1.

However, simply modeling CDS as a put option on the asset fails to generate appreciable losses given default, and therefore fails to generate empirically plausible levels of the CDS spread. Banks are highly levered, and therefore have assets relative to liabilities ($A^i_t$) not too far above 1. Given that defaults are rare, the Brownian component of $A^i_t$ cannot have too high of a volatility. An appreciable loss, i.e., the event that $A^i_{t+\tau_D} \ll 1$, is therefore a large number of standard deviations away from the mean, and therefore effectively a probability zero event. Only jump risk can generate large shortfalls, but only if the jumps are assumed to be large in the first place ($V$ has high probability of being small. A Merton-style model therefore effectively equates CDS to large jumps, and almost completely ignores any other type of default.

Instead, I use a fixed recovery rate $\alpha \in (0, 1)$, as is common both in the literature and industry models. In the context of the Merton model, this can be thought of as including an additional cost to default. The standard values for senior debt to US financial firms is $\alpha = 0.4$ (a loss-given-default value of 60%). Assuming that the incidence of a default has a constant hazard rate, this leads to

\[ (A^i_{t+\tau_D} - 1)^+ - k^+ = (A^i_{t+\tau_D} - (1 + k))^+. \]

\[ (k - (A^i_{t+\tau_D} - 1))^+ = ((1 + k) - A^i_{t+\tau_D})^+ - (1 - A^i_{t+\tau_D})). \]

For academic papers, examples include Duffie and Singleton (1999); Veronesi and Zingales (2010). For industry models, an example is the ISDA model, which is outlined in White (2013).
a model-implied CDS spread of
\[ \text{CDS}_t^{[i]} = -\frac{1 - \alpha}{\tau D} \log \Pr_t^{[i]} [A_{t+\tau D} > 1] \]

Explicit formulas for each of the asset prices are given in Appendix E.1.

4.2 Objective Function for Jump Risk

Let \( S_t^{[i]} = (A_t^{[i]}, \sigma_t^{[i]}, \lambda_t^{[i]}) \) denote the unobserved state variable values at time \( t \) for bank \( i \), and let \( \Omega_t^{[i]} = (\mu_t^{[i]}, \sigma_t^{[i]}) \) denote the parameters of the model. Time \( t = 1, 2, \ldots \) is indexed in trading days, and let \( T \) be the total number of observations. The objective is to solve the following minimization problem

\[
\min_{(\Omega_t^{[i]}, S_t^{[i]})_t} \sum_{t=1}^{T} \left( \sum_{j=1}^{M_t^{[i]}} \left( \frac{\kappa_{t,j}^{[i]} - \tilde{\kappa}_{t,j}^{[i]}(\Omega_t^{[i]}, S_t^{[i]})}{M_t^{[i]}} \right)^2 + \frac{\hat{\eta}_t^{[i]} k}{M_t^{[i]}} \left( \frac{\sigma_E(\Omega_t^{[i]}, S_t^{[i]}) - \sqrt{\hat{v}_t^{[i]}}}{\hat{v}_t^{[i]}/2} \right)^2 \right)
\]

s.t. \( \text{CDS}_t^{[i]} = \text{CDS}(\Omega_t^{[i]}, S_t^{[i]}) \) \( t = 1, \ldots, T \)
\( E_t^{[i]} = E(\Omega_t^{[i]}|\text{the}, S_t^{[i]}) \) \( t = 1, \ldots, T \)

Notation in this objective is as follows. \( \kappa_{t,j} \) is the Black-Scholes implied volatility (BSIV) of equity option \( j \) at time \( t \). \( \tilde{\kappa}_{t,j}^{[i]}(\cdot, \cdot) \) is the model-implied BSIV of an option with the same strike as option \( j \) at time \( t \). \( M_t^{[i]} \) is the number of options at time \( t \). \( \hat{\eta}_t^{[i]} \) is a consistent estimate of the first term (the mean sum of square deviations between model-implied and measured BSIVs at time \( t \)), and is used for scaling purposes. \( \sigma_E(\cdot, \cdot) \) is the model-implied equity volatility. \( \hat{v}_t^{[i]} \) is an estimate of equity variation using \( k \) intraday price increments. Details of this estimator in my context are explained in Section 4.3 below. \( \text{CDS}_t^{[i]} \) is the CDS spread in the data at time \( t \), and \( \text{CDS}(\cdot, \cdot) \) is the model-implied CDS spread given the parameters. Lastly, \( E_t^{[i]} \) is the value of equity (relative to liabilities) at time \( t \), and \( E(\cdot, \cdot) \) is the value of equity relative to liabilities in the model.

Although it involves a lot of notation, the estimator itself is relatively straightforward. The first term is a non-linear least squares objective in units of BSIV. Observations within each day are equally weighted, and observations across days are weighted by the relative number of options available at each day. The second term is a penalty term that penalizes large deviations between the model-implied equity volatility and a measure of realized volatility. The scaling term \( k/M_t^{[i]} \) captures the relative amount of data used to compute each term. The scaling term \( \hat{\eta}_t^{[i]} \) increases the penalty when the model performs poorly, and decreases the penalty when the model fits the data well.

The \( \Omega_t^{[i]} \) portion of the minimization problem of equation (10) is placed in parentheses to highlight that the parameters are held fixed for much of the empirical specification. The estimator
has a tendency to set $\mu_J \approx 0$ and $\sigma_J$ large in order to generate a mixture of small and large volatility states of the world. To isolate out large negative jumps, I instead set $\mu_J = -0.5$ and $\sigma_J = 0.9$. Only for aggregate jump risk, which I discuss below, do I jointly estimate the state variables and parameters.

The objective in (10) features two changes relative to Andersen et al. (2016): 1) the new state variable $A_t^i$, and 2) the two constraints for equity and CDS. Andersen et al. (2016) is performed in an unlevered environment, and the process being estimated is that of the stock price, not an underlying latent asset. The stock price (their equivalent to $A_t^i$) is observable, and therefore does not need to be estimated. To discipline the estimator’s choice of $A_t^i$, I include equity as a constraint. I include the CDS spread as a constraint rather than as another option-like security with a Black-Scholes implied volatility for the simple reason that this paper is primarily concerned with decomposing default risk, and therefore the estimator should provide a decomposition of the default risk. Having CDS as a constraint forces the estimator to match the empirical equivalent of default risk.

To solve the minimization problem in (10), I use the IPOPT optimizer. To obtain estimates of $\hat{\eta}_t^i$, I first run the minimization with $\hat{\eta}_t^i$ set to the variance of the BSIVs in the data for that day. I use the resulting estimates of $\Omega^i$ and $S_t$ to estimate $\hat{\eta}_t^i$, and run the estimation procedure again to obtain my estimates for the parameters and state variables.

Given solution to the problem in (10), I construct the default probabilities as follows. The Brownian default probability at bank $i$ at time $t$ ($\text{Brownian}_t^i$) is calculated as the probability of default with the jump intensity set to zero. The jump default probability at bank $i$ at time $t$ ($\text{jump_like}_t^i$) is calculated as the residual.

### 4.3 Intraday Variation ($\hat{\nu}_t^i$)

The intraday variation estimate $\hat{\nu}_t^i$ is based on the high-frequency volatility estimator in Andersen et al. (2015). Time is still measured in days, but can now take on fractional values. Stock prices are measured at $n$ intervals throughout the day. The value $P_t^i$ is the end-of-day price of the stock, and $P_{t-j/n}^i$ denotes the price $j$ time intervals before the end of day. Let $\Delta_j P_t^i = \log P_{t-j/n}^i - \log P_{t-(j+1)/n}^i$ denote the price increment $j$ intervals before the end of the day. I start by computing the value

$$\hat{w}_t^i = \frac{n}{k} \sum_{j=-k+1}^{0} (\Delta_j P_t^i)^2 1 \left\{ |\Delta_j P_t^i| \leq \alpha n^{-\varpi} \right\}$$

Note that there are two modifications to the usual total variation estimator. First, $k$ may be less than $n$, in which case we use a smaller sample of price increments. Andersen et al. (2015) use this

---

9 As long as $\mu_J \ll 0$, the choice of $\mu_J$ is inconsequential; a jump leads to default with probability close to 1.

10 IPOPT is part of the COIN-OR project, is available at https://projects.coin-or.org/Ipopt, and is based on Wächter and Biegler (2006). I use the linear solvers in the HSL library.
in proofs to obtain a closer estimate to the instantaneous volatility at the end of the day. I set \( k \) close to \( n \) in such a way to discard the first 5 minutes of the trading day. Second, the estimate throws out particularly large price differences. I follow Andersen et al. (2015) in the choice of \( \alpha \) based on the last day’s variation and \( \varpi = 0.49 \).

Andersen et al. (2016) consider very short-term options, and are therefore satisfied with this estimate of variation. In my context, however, it is inappropriate. I have a fixed \( \sigma_{t}^{[i]} \) in the approximate bank asset dynamics equation (9), yet I will be considering options at longer horizons (\( \tau^D \) is six months). To address this, instead of plugging in the spot volatility (which is estimated by \( \hat{w}_{t}^{[i]} \)), it is more appropriate to plug in an estimate for the average volatility over the next \( \tau^D \) units of time.

To obtain the smoothed forecast, I fit an AR(1) process

\[
\hat{w}_{t+1}^{[i]} = \rho(\hat{w}_{t}^{[i]} - \bar{w}) + \epsilon_{t+1}
\]

using OLS, and then set

\[
\hat{v}_{t+1}^{[i]} = \frac{1}{\tau^D} \sum_{j=0}^{\tau^D-1} \mathbb{E}[\hat{w}_{t+j}^{[i]} | \hat{w}_{t}^{[i]}] = \frac{1}{\tau^D} \sum_{j=0}^{\tau^D-1} \left( \bar{w} + \rho^j (\hat{w}_{t}^{[i]} - \bar{w}) \right)
\]

I then rescale the units to annualized volatility.

### 4.4 Aggregate Jump

To compute an estimate of the aggregate jump intensity \( \lambda_t \), I apply an unlevered version of the same methodology to the S&P500. The processes in equations (8) and (9) are now matched to the observable level of the S&P500, and not an unobserved latent asset. Options are now treated as directly on the asset level, and not on the residual. The estimator (10) is adjusted accordingly, and the two constraints are removed. The estimation is identical to Andersen et al. (2016) except for the smoothed volatility \( \hat{v}_t \).

From the estimated jump intensities, I compute the aggregate jump probability as the probability of at least one jump occurring during the next \( \tau^D \) units of time.

### 4.5 Data

Data for this project come from multiple sources, and involve a fair amount of processing before they can be plugged into the estimator. I only present a brief summary here; the full details of the data are given in the Data Appendix A.

Six month CDS spreads are obtained for senior debt on the holding company for each of the firms from Markit. I use the modified restructuring clause, as these quotes had the highest quality
ratings in the time periods used.

Intraday equity prices used to compute the intraday variation are computed using the NYSE Trade and Quote (TAQ) data. I use the trade data to compute the last traded price before the start of each minute. For each day, I use data from 9:35 to 16:00. I remove half-days from the sample.

Options prices are obtained from OptionMetrics. For each of the tickers, I choose options with maturity closest to six months to match the CDS data. I then drop quotes with a bid-ask spread more than four times the bid. I compute the Black-Scholes Implied Volatilities for each of the options, using imputed LIBOR rates available on FRED to get the a rate for the same maturity as the option. I drop all options with moneyness greater than 3, as my estimation strategy is not meant to deal with large positive jumps. I also drop the deepest out-of-the-money options until the midpoint price for the most out-of-the-money option is less than all the other prices to prevent clear sources of arbitrage.

I use CRSP to obtain market capitalizations for each of the banks. I scale this by a quarterly estimate of liabilities, obtained from FR Y-9C data for bank holding companies and from Compustat for other firms.

Finally, for the aggregate jump intensities, I use two different types of data on the S&P500. I use S&P500 options (SPX), which are written directly on the S&P500, and apply the same cleaning methodology as used for the banks. For intraday prices, as well as daily levels, I use prices on the SPDR S&P500 ETF (SPY) which closely tracks the S&P500.

5 Results

5.1 Ridge Regression

Regression equation (7) suffers from being poorly conditioned. The right-hand-side variables are moderately correlated, and once considered all together, the number of parameters estimated is a sizable fraction of the number of data points available.

To overcome the high variance of the point estimates, I use a form of ridge regression adapted to the problem of estimating a network. Ridge regression, developed in Hoerl and Kennard (1970), has the ability to reduce the mean square error of parameter estimates by introducing bias to the coefficients. The objective of minimizing the sum of square errors is augmented by the $L^2$-norm of the parameters multiplied by a penalty parameter. This is mathematically equivalent to adding orthogonal noise to each of the regressors. For an appropriate choice of the penalty parameter, the reduction in variance of the estimator outweighs the bias when measured as mean squared error.

The standard issue with using ridge regression is the choice of the ridge parameter. Hoerl et
al. (1975) derive an optimal parameter for when the regressors are scaled to correlation form\textsuperscript{11} and are orthogonal to each other,\textsuperscript{12} and suggest that this parameter remains near-optimal when the regressors are not orthogonal. This optimal choice of the penalty parameter depends on the true parameters, which much of the subsequent literature has worked to address estimating.

The choice of scaling the regressors to have equal variance is not necessarily appropriate when the magnitudes of the unscaled coefficients are in equal units. Consider the case where there are two regressors, $x_1$ and $x_2$, with $\text{var}(x_1) = 10$ and $\text{var}(x_2) = 1$. When rescaled, the coefficient on $x_1$ is multiplied by a factor of $\sqrt{10}$. Minimizing mean square error of the scaled coefficients therefore tilts toward minimizing the mean square error of the coefficient on $x_1$. A factor of 10 times as much mean square error on the coefficient on $x_2$ is tolerated for reducing the mean square error on $x_1$.

In Van Vliet (2018), I develop a new method for generating a ridge parameter that is adapted to the problem considered here. I compute the optimal ridge penalty parameter without rescaling the regressors, under the assumptions about the magnitude of the coefficients. Specifically, I assume some sparsity in the coefficients (modeled as a coefficients being non-zero with some probability) and I assume that the expected magnitude of the coefficient if it is non-zero is known.\textsuperscript{13} Under these assumptions about the ex-ante distribution of parameters, I compute an optimal ridge penalty that minimizes the expected mean square error given the values of the regressors.

### 5.2 2008 Financial Crisis

As a first application, I run my methodology using the largest banks, as well as a select number of other large financial institutions, around the financial crisis. Firms are selected on the basis of appearing to have a liquid market for their CDS, as reflected by frequent updates in the quotes available. All estimates were performed using ridge regression (as outlined above), with the optimal ridge parameter set under the assumption that nonzero coefficients have a value of around 1 and that each coefficient has a 30\% chance of being nonzero.

[Table 3 about here.]

[Table 4 about here.]

[Figure 1 about here.]

**Bear Stearns Episode**  Table 3 shows the results using the two years prior to March 13, 2008, the last day before the collapse and sale of Bear Stearns. Bear Stearns, Lehman Brothers, and to

\textsuperscript{11}In the usual setup of stacking rows of the right hand side variables into a matrix $X$, this means that the variables are scaled so that $X'X$ has unit diagonal.

\textsuperscript{12}In the notation of the previous footnote, this means $X'X = I$.

\textsuperscript{13}Technically, I only make assumptions about the expectation of each coefficient squared and the expectation of the product of two coefficients.
some extent Capital One and Goldman Sachs are all estimated to have the potential for devastating spillovers upon default. Interestingly, the implied connections appear to be quite broad. An increase in Brownian risk is associated with an increase in jump risk at most of the remaining banks. This type of connection would be consistent with either 1) these banks being highly connected, or 2) a default at these banks leading to contagion in the entire financial sector.

Table 4 shows the same regressions with just two more trading days added. The noticeable difference is that Bear Stearns no longer appears to be connected, despite only two having to more data points. This is entirely consistent with the orderly way in which Bear Stearns collapsed. On Friday, March 14, 2008 the Federal Reserve announced its approval of the financing arrangement to sell Bear Stearns to JPMorgan Chase.14 By Sunday, March 16, 2008, the Federal Reserve had announced the Primary Dealer Credit Facility, lowered the primary credit rate, and raised the maximum maturity of loans.15 The two actions were intended to bail out Bear Stearns and to “promote orderly market function.”16 Although the counterfactual is of course unknown, the data appears to support that Bear Stearns’s failure had minimal impact. Figure 1 plots both the cumulative equity returns and the CDS-implied default probabilities around the event. With the exception of Lehman Brothers, little appears to happen to other financial institutions. This lack of response from other financial institutions to the large increase in default probability at Bear Stearns drives the change in estimates in Table 4.

[Table 5 about here.]

[Table 6 about here.]

[Figure 2 about here.]

[Table 7 about here.]

**Lehman Brothers Episode** Table 5 shows the results up to two weeks before the failure of Lehman Brothers. The estimator now highlights the potential for Bank of America, Goldman Sachs, and Lehman brothers to generate spillovers should one of them default. Again, the implied connections are broad. An increase in Brownian risk is associated with an increase in jump risk at most of the remaining banks.

Table 6 shows the same regressions right up to the default of Lehman Brothers. As with the Bear Stearns episode, the coefficients on Lehman Brothers vanish when we include data up to the last trading day. This is consistent with the ex-post narrative of the lack of Lehman’s connectedness. Though the losses to creditors were substantial, the losses were spread out and not concentrated at

---


16Ibid.
any particular lender (Scott, 2016, Chapter 4). Furthermore, despite fears about the large notional value of CDS contracts written on Lehman bonds requiring large payouts, the net payments were relatively small (Scott, 2016, Chapter 4.2). Figure 2 shows the final days of Lehman, and confirms that although the default lead to equity losses and eventual increases in implied-default probability, the changes are not extreme.

Interestingly, neither Table 5 nor Table 6 indicate that AIG was the source of any connections. Table 7 shows the results with the two additional trading days leading up to AIG’s bailout included. Despite the implied default probability rising rapidly and the market value of AIG falling during these last days (see Figure 2), the estimated coefficients are trivial. This may reflect the market belief that a failure at AIG would be met with a bailout. This may also reflect that no particular counterparty of AIG would have been subject to substantial losses had AIG defaulted (Scott, 2016, Chapter 4.2).

End of 2008 As a last exercise with data from the 2008 Financial Crisis, I consider data up to the end of 2008. Table 8 shows the results. The two notable changes are the increase in the coefficients for JPMorgan Chase, and the decrease in the coefficients for Goldman Sachs. JPMorgan by then had absorbed not only Bear Stearns but also Washington Mutual, which was at the time the largest savings and loan institution, increasing the market evaluation of its risk. Meanwhile, Goldman Sachs and Morgan Stanley had become bank holding companies, thereby gaining access to additional government protection.

5.3 Current Estimates

Table 9 considers data for the two years prior to the end of 2017, the most recently available data. All the estimated coefficients, besides loadings on the aggregate jump, are zero. This is a direct result of the ridge regularization technique used. Brownian default risk, both in levels and first differences, is orders of magnitude smaller than total default risk during this period. Any signal, even if present, is simply too small and is overwhelmed by the noise.

This result, or rather the lack thereof, highlights an important aspect of the estimation strategy. Although it relies on prices being forward looking, and hence relies on agents being extrapolative, the estimation strategy itself is not extrapolative. Market participants must have real fear of a default and its spillovers before it is priced into securities and can be detected using the methodology

---

presented. Given the high level of capitalization at banks post-crisis, a Brownian default has almost zero probability.\footnote{During this time sample, a Brownian default would require at least 5.5 standard deviations of negative shocks, with a mean above 9.}

6 Robustness

Assumption 1 and its weaker alternative Assumption 2 both rely on idiosyncratic jump risk being sufficiently uncorrelated with Brownian risk at each of the banks. I now discuss the cases where these assumptions may fail to hold. In each case, I argue that although it is impossible to rule out violations of the identification assumptions, the controls present in regression (7) mean that we need an unlikely combination of events for violations to lead to substantial spurious coefficients.

6.1 Violations of the Identification Assumptions

Broadly speaking, there are three ways these restrictions can fail: correlation, reverse causation, and measurement.

6.1.1 Correlation

Correlation can be thought of as an omitted variable: a driver of jump risk at bank \(i\) that also drives Brownian risk at bank \(j\). The inability to control for this omitted variable would lead to non-zero \(\gamma_{ij}\) values, even if there is no causal relationship between the two banks. Two controls present in regression equation (7) help absorb the influence of this type of omitted variable: the control for aggregate jump risk and the control for the bank’s own Brownian risk. I argue below that these two controls, combined with attenuation bias, make it highly unlikely that omitted variables lead to large, spurious coefficients.

Controlling for aggregate jump risk means that we are looking for a common driver that drives Brownian risk at one bank and idiosyncratic jump risk at another, but is not captured by aggregate jump risk. This effectively rules out aggregate factors from generating spurious results. We would need an aggregate factor that is a large contributor to jump risk as at the bank, but that contributes very little jump risk to the market as a whole, almost contradicting that it is an aggregate factor in the first place.

The control for the bank’s own Brownian risk further reduces the possibility that an aggregate or idiosyncratic factor drives correlation. Assumption 2 allows for a common driver of both bank \(i\)’s and bank \(j\)’s Brownian risk to correlate with bank \(i\)’s idiosyncratic jump risk. It is only the \(residual\) of bank \(i\)’s idiosyncratic jump risk after being regressed on aggregate jump risk and bank \(i\)’s Brownian risk needs to be uncorrelated with bank \(j\)’s Brownian risk. Therefore, this common
driver would have to affect the Brownian default risk of $j$ and the idiosyncratic jump risk of $i$, while not affecting the Brownian risk of $i$, in order to generate spurious results. Such a driver cannot be ruled out entirely, as in the tranched example I consider below, but is unlikely.

Finally, both banks would have to be heavily driven by their common factor. Letting bank $j$ be the bank with the Brownian exposure while bank $i$ has the jump exposure, to prevent substantial attenuation bias, variation in bank $j$’s Brownian default risk would need to be primarily driven by the omitted variable. With bank $j$’s Brownian risk having a substantial exposure to this factor, variation in bank $i$’s jump risk would then need to be correspondingly large to obtain a substantial coefficient. The need for both banks to be heavily exposed to the common factor again makes spurious correlations unlikely.

While the three conditions above are already unlikely when considered individually, all three are needed together to generate large spurious coefficients, making spurious, large coefficient estimates even less probable. This limits the scope for spurious results to common drivers that are 1) idiosyncratic in the sense that they are not reflected in aggregate jump risk, 2) a large portion of both banks’ balance sheets, and 3) found in the jump risk of one bank and found in the Brownian portion of only the other bank.

**Example: A Common Project**  As an example, suppose bank 1 and bank 2 are both invested in a project, and suppose this project has both Brownian and jump risk that are correlated in the sense that negative Brownian shocks tend to correlate with increases in the expected jump intensity. Then, both banks would have idiosyncratic jump risks that are correlated with their own, and each other’s, Brownian risks. Including each bank’s own Brownian risk is therefore likely to capture a portion of the correlation, and drive out the spurious coefficient.

In this same example, having one of the banks hedge against their jump risk would not suffice to generate spurious results. The unhedged bank would still have the Brownian risk of the project as well. To get spurious results, one of the two banks would effectively have to hedge against Brownian risk, and maintain only the jump risk of the project. This is possible through carefully crafted tranching. If we assume that jumps tend to lead to larger losses than bad Brownian draws, and then have one bank take the senior tranche while the other takes the junior tranche, the Brownian risk of the bank with the junior tranche would be correlated with the jump risk of the bank with the senior tranche. For this to turn into large coefficients, we would then need both banks to hold large amounts of the tranched risk. If the bank with the senior tranche has only a small exposure to the junior tranche, other risk is likely to dominate. Likewise, if the bank with senior tranche has only a small exposure to the senior tranche, estimated coefficients would be small.
6.1.2 Reverse Causality

Reverse causality is statistically similar to correlation, but has a different interpretation. In this case, we would need that bank \( j \) is actually exposed to bank \( i \), but that this connection is reflected in \( i \)'s jump risk correlating with bank \( j \)'s Brownian risk. As with correlation, the control for the bank’s own Brownian risk and attenuation bias make it unlikely to lead to large, results. Furthermore, reverse causality requires transforming jump risk into an inherently smoother Brownian risk. Besides the case of mismatched maturities, which can smooth out jump risk into Brownian risk through expectations about future jump risk, it is difficult to come up with connections that result in reverse causality.

Consider the two bank case for simplicity. Suppose that bank 2 has exposure to bank 1, and that this exposure is through the Brownian part. To a first-order expansion, we have

\[
\Delta \text{Brownian}_{2t+1} = \Delta \text{Brownian}_{\text{own}}_{2t+1} + \alpha \Delta \text{jump}_{\text{like}}_{1t+1}
\]

where \( \Delta \text{Brownian}_{\text{own}}_{2t+1} \) is the portion of the Brownian risk at bank 2 unrelated to the exposure. The population coefficient for regression equation (7) is

\[
\gamma^1_2 \equiv \frac{\text{cov}(\Delta \text{jump}_{\text{like}}_{1t+1}, \Delta \text{Brownian}_{2t+1})}{\text{var}(\Delta \text{Brownian}_{2t+1})} = \alpha^{-1} \left( 1 + \frac{\text{var}(\Delta \text{Brownian}_{\text{own}}_{2t+1})}{\alpha^2 \text{var}(\Delta \text{idio}_{\text{jump}}_{1t+1})} \right)^{-1}
\]

where here the tildes indicate the residuals after regressing on \( \Delta \text{agg}_{\text{jump}}_{t+1} \) and \( \Delta \text{Brownian}_{1t+1} \). Here, the first term \( (\alpha^{-1}) \) is the reverse causality, and the second term is the attenuation bias generated by orthogonal movements in the Brownian risk at bank 2. The derivation can be found in Appendix E.3.

Equation (12) highlights the two features we need for reverse causality to lead to large coefficients. First, we need \( \alpha \) to be small \( (\alpha \gg 1) \), as otherwise large changes in Brownian risk at bank 2 are only associated with small changes in jump risk at bank 1. Second, we need bank 2 to have little own Brownian risk. In other words, bank 2 needs to be very heavily invested in bank 1, with little other exposure. Otherwise, the usual attenuation bias will shrink the estimated coefficient.

As with the pure correlation case, the controls help mitigate the impact of these variables. The tildes in equation (12) indicate controlling for the Brownian default risk at bank 1, as well as the aggregate jump risk. If jump risk at bank 1 leads to Brownian risk at bank 2, it is likely to be reflected in the Brownian risk of bank 1 as well. This control therefore reduces the denominator in the second term of (12), further suppressing the overall coefficient. As before with the correlation case, this does not rule out spurious results, but it makes large, spurious results unlikely.
**Example: Long Maturity Debt** To get a more concrete sense of what is required to get a large coefficient, consider the case where bank 2 is invested in a long-term debt claim to bank 1. If the debt has maturity $\tau > 0$ less than the time horizon $\tau^D$ of the analysis, there would not be a problem. The jumps that make up jump risk at bank 1 would be realized before $\tau^D$, and hence would also be realized at bank 2 before $\tau^D$ units of time have passed. Brownian risk at bank 2 would not be influenced by these jumps. Issues arise only when the maturity $\tau$ is longer than $\tau^D$, in which case jump risk between $\tau^D$ and $\tau$ can present itself as Brownian risk.

When looking at a horizon $\tau^D$ that is shorter than the maturity of the debt claim $\tau$, the value of assets at bank 2 are exposed to the holding period return of the debt at bank 1, instead of the fully realized return of the debt at maturity. This holding period return is influenced by three factors: 1) any defaults that happen before $\tau^D$, 2) any innovations in the Brownian portion of bank 1’s assets before $\tau^D$, and 3) any innovations in the expected jump risk of bank 1’s assets between time $\tau^D$ and $\tau$. This third factor can transform information about jump risks into Brownian risk, and can therefore lead to reverse causality.

Three things must hold to obtain substantial reverse causality. First, future jump risk at bank 1 must be sufficiently unpredictable. Levels of the future jump risk already known when the debt is purchased do not matter. Time 0 expectations about jump risk between time $\tau^D$ and $\tau$ is already priced into the debt at time 0, and does not contribute to a negative holding period return. Only changes between time 0 and time $\tau^D$ of the expected future jump risk leads to movements in the holding period return; large increases in the expected jump risk lead to large negative holding period returns. Therefore, for future jump risk to lead to Brownian default risk at a time horizon $\tau^D$, innovations in expectations about future jump risk must have sufficiently high variance. Furthermore, these innovations in the expectations must be Brownian in nature. Jumps in these expectations would be jumps reflected as jumps in the debt price, and would show up in the jump risk of bank 2.

Second, information about future jump risk must be correlated with information about current jump risk. Bank 1’s jump risk is measured as the jump risk between time 0 and time $\tau^D$, and not the jump risk from $\tau^D$ to time $\tau$ that is reflected in the debt holding period return. To get reverse causality, we need that when current jump risk is higher, the chance of drawing a sufficiently large future jump risk is also be higher. Of the three conditions, this one is likely to hold, as this is essentially just the statement that jump risk is autocorrelated.

Third, bank 2 must be heavily invested in bank 1. Something as large as a 10% increase in expectations about future jump default risk would lead to at most a 10% reduction in the value of a debt claim. Bank 2 would need to have a very large percentage of its assets be this debt claim, or be very close to defaulting already, in order for this to substantially alter the probability of default. If bank 2 has diversified holdings, large changes in bank 1’s jump default probability are likely only to lead to small changes in bank 2’s default probability, and the estimated coefficient would be suppressed by attenuation bias.
Even if all three conditions were to hold, it is still unlikely to lead to large coefficients being estimated. In order for jump risk at bank 1 to turn into Brownian risk at bank 2, movements in jump risk have to be large enough to affect the debt price. Unless bank 1 is highly capitalized, or somehow insulated from defaulting at an intermediate time due to the low value of its assets, such movements should also depress the value of bank 1 sufficiently to trigger a default. This would be reflected in the Brownian risk at bank 1 as well, and therefore be partially taken out in the regression context.

Given all of the necessary conditions, it therefore seems highly unlikely that holding longer maturity debt claims would lead to a substantial reverse causality. The large movements in jump risk would only result in small movements in the Brownian component, which are unlikely to trigger default. If they lead to large enough movements, this is also likely to trigger a Brownian default in the originating bank as well.

6.1.3 Mismeasurement

Lastly, we need to consider mismeasurement. The discussion so far has assumed that I can perfectly isolate default risk due to jump-like behavior and default risk due to Brownian shocks. Now I briefly turn to what happens if that decomposition is imperfect. I again argue that mismeasurement is unlikely to generate spurious coefficients.

**Brownian Risk as Jump Risk (Fat Tails)** The approximating model in equation (9) fixes volatility, and therefore the diffusion portion of the model generates log-normal distributions. A typical stochastic volatility setup, on the other hand, will tend to generate fatter tails. If the true diffusion process has more mass in the tails, the model may attribute too little default probability to the Brownian piece.

Although this affects the magnitudes of the coefficients estimated, it is unlikely to cause spurious non-zero coefficients. As long as the mismeasurement is a sufficiently smooth function, the contribution to the measured jump risk is approximately linear in the measured Brownian risk, and the measured Brownian risk is approximately a scaled version of the true Brownian risk. The estimates for $\gamma_j^{[i]}$ in regression (7) are then scaled versions of the true coefficients, much as in Proposition 1. A non-zero value still indicates the presence of a linkage. Furthermore, the scaling of the coefficients is moderate. Misattributing a fraction $\alpha$ of the Brownian risk as jump risk scales the coefficients by $(1 - \alpha)^{-1}$. As long as the majority of the default probability is not falsely measured as jump risk, the coefficients will remain the same order of magnitude.

**Jump Risk as Brownian Risk** The opposite case has the potential to generate spurious coefficients. Again, assuming that the mismeasurement is a smooth function of measured jump risk, the measured Brownian risk is approximately linear in the measured jump risk, and the measured jump
risk is approximately a scaled version of the true jump risk. This causes condition (i) in Assumption 2 to fail. In the notation used in this assumption, $\Delta \text{Brownian}_{i,t+1}$ may now contain information about $\Delta \text{idio\_jump}_{j,t+1}$ through idiosyncratic jump risk at $j$ spilling over to $i$. The coefficient $\gamma_{ij}$ may therefore be non-zero despite the causal relationship being from $j$ to $i$.

However, of the two cases, this one is less likely. Given the simple model with log-normal Brownian distributions, and the substantial evidence that financial returns are fat-tailed, the estimator is much more likely to misattribute Brownian risk as jump risk rather than the opposite.

### 6.2 Empirical Test of Reverse Causality

Assumptions 1 and 2 are not directly testable as the restrictions pertain to unobservable portions of the default probability. However, I can test these restrictions in an exaggerated setup. I assume that each bank is only exposed to jump risk, and that this jump risk has a purely Brownian evolution. I then construct a hypothetical bank that is a levered debt claim on the bank in question, and has no other assets. I form this debt claim by constructing a synthetic bond from CDS, and compute its price dynamics by fitting a GARCH time series model to past synthetic returns.

For this exercise, let $t$ be indexed in days. For each bank $i$, let $\text{CDS}_{6m,i,t}$ denote the 6-month CDS spread on day $t$, and let $\text{CDS}_{1y,i,t}$ denotes the 1-year CDS spread on day $t$. In the absence of frictions, the yield on a zero-coupon debt claim equals the risk-free rate plus the CDS spread of the same maturity. Therefore, the 6-month holding period return for a synthetic 1-year debt claim is

$$\log \text{HPR}_{6m,i,t} = \frac{1}{2} \left( \text{CDS}_{6m,i,t} + \text{CDS}_{1y,i,t} - r_{6m,t} - r_{1y,t} \right)$$

where $r_{6m,t}$ and $r_{1y,t}$ are the 6-month and 1-year risk-free rates, respectively.

Now, consider a hypothetical bank that is a levered synthetic debt claim on bank on bank $i$. Suppose it starts with $A_t$ assets against liabilities with a present value of 1. For simplicity, assume its liabilities also mature in one year. Over the six months, log assets accumulate according to the holding period return in (13), while log liabilities grow from 0 to $r_{1y,t} - \frac{1}{2} r_{6m,t+6m}$. The hypothetical banks therefore defaults if the excess holding period return falls below its asset buffer

$$\log \text{eHPR}_{6m,i,t} = \text{CDS}_{1y,i,t} - 0.5 \text{CDS}_{6m,i,t+6m} \leq - \log A_t$$

If we assume that all innovations in CDS spreads are Brownian, and not themselves a jump process, the Brownian default risk for the hypothetical bank is the probability of the condition in equation (14) occurring.

To estimate the probability of a sufficiently negative holding period return, I fit an MA(125)-GJR-GARCH(1,1) model to the data. The MA(125) model for the mean captures the six month overlap in realizations of holding period returns. For computational tractability, I set all of the moving average coefficients to 1. The GJR-GARCH(1,1) model for the error variance of Glosten
et al. (1993) is a modification of the usual GARCH model that allows for asymmetric responses to positive and negative shocks. Details, including the specification of the model, and how I use simulation to compute Brownian default probabilities, are in Appendix G. For the asset level log $A_t$, I consider several different values corresponding to different levels of capitalization.

Given the simulated Brownian default probabilities at the hypothetical bank, I consider the following falsification regression

\begin{equation}
\Delta \text{jump-like}_{t+1}^{[i]} = \alpha^{[i]} + \beta^{[i]} \Delta \text{Brownian}_{t+1}^{[i]} + u_{t+1}^{[i]}
\end{equation}

Here, $\text{jump-like}_{t+1}^{[i]} = 1 - \exp\{-\text{CDS}_{6m,t}^{[i]} \tau^D / (1 - \alpha)\}$ is the implied jump default risk priced into CDS assuming all defaults are caused by a jump. As before, $\alpha = 0.4$ is the assumed recovery, while $\tau^D = 0.5$ reflects the six month window. $\text{Brownian}_{t+1}^{[i]}$ is the Brownian risk of the hypothetical bank invested solely in bank $i$.

[The implementation will be presented here. I temporarily removed it to fix a coding error.]

7 Conclusion and Further Research

This paper develops a new methodology for discovering connections between financial institutions from market price data. The estimator uses information embedded in the prices of equity, equity options, and credit default swaps to separate default risk into the portion attributable to jumps and the portion attributable to smoother shocks. Connections that appear as large spillovers upon default are contained entirely in the jump risk portion, allowing for connections to be estimated by regressing jump risk on non-jump risk. An application to the 2008 Financial Crisis shows that this is a promising novel approach to estimating connectedness in future financial downturns. It also shows that market beliefs about connections appear to be broad; if a financial firm has spillovers on one firm, it tends to have spillovers to many firms. This suggests that market participants are concerned about a high level of interconnectedness, or that they anticipate broad contagion upon a default.

One potential modification that may be fruitful for further research is to extend the method used to price securities to consider multiple maturities. The model used to isolate jump risk treats a firm as a levered asset with a single maturity, thereby requiring that both the options and the credit default swaps used are of the same maturity. As a result, this paper uses quotes for CDS spreads at the shortest available time horizon: six months. A richer model that includes an intermediate period could allow for using both the shorter time horizon options data with longer time horizon
CDS data, thereby using the more liquid versions of both.\footnote{At a long time horizon, such a model would need to have several features that are unnecessary in the short-run. For example, the log-normal distributions arising from a fixed volatility become less plausible.}

Another avenue for further research is to consider a more extrapolative approach to estimating spillovers. The estimation methodology employed in this paper is fundamentally not extrapolative. Market participants perform any necessary forecasts about potential defaults, and then reflect this in asset prices by placing higher probabilities on default states of the world. The benefit is that I am not relying heavily on a model to draw conclusions about crisis states of the world from correlations obtained in non-crisis states. The drawback is that this approach cannot forecast problematic connections when financial institutions are well capitalized, as in the example using data from 2016-2017. Spillovers from non-jump defaults become too small relative to other jump-like defaults to allow for connections to be estimated. Perhaps a more extrapolative version of the approach, where jumps that do not necessarily lead to default are considered, would enable the methodology to forecast problematic connections further in advance.

Finally, it is worth noting that the setup can be used to analyze other types of firms, or even connections at different scales. At no point does the setup require that the assets are for financial firms. The methodology can be used to analyze any sufficiently rich set of options and liquid debt-like instruments. Furthermore, with some minor modifications, the methodology can be applied to sectors or countries instead of individual firms.

References


HSL, “A Collection of Fortran Codes for Large Scale Scientific Computation.” http://www.hsl.rl.ac.uk/.


Cumulative returns and CDS-implied 6 month default probabilities around the collapse of Bear Stearns (BSC) in 2008. The Federal Reserve approved a financing arrangement for JPMorgan to purchase Bear Stearns on the morning of Friday, March 14, 2008.

Figure 1: Final Days of Bear Stearns.
Figure 2: Final Days of Lehman Brothers.

Cumulative returns and CDS-implied 6 month default probabilities around the bankruptcy of Lehman Brothers (LEH) in 2008. On September 15, 2008, Lehman Brothers filed for Chapter 11 bankruptcy protection.
Table 1: Notation and definitions for types of defaults.

All default probabilities are defined over a time horizon of \( \tau^D > 0 \). Here, \( i \) denotes a bank, and \( t \) denotes the current time. \( D_{t-}^i \) is a indicator that bank \( i \) has not yet failed prior to time \( t \), i.e., \( A_s^i > 1 \) for all \( s < t \). For convenience, define \( dN_t = N_t - \lim_{s \uparrow t} N_s \), and \( dN_t^i = N_t^i - \lim_{s \uparrow t} N_s^i \).

<table>
<thead>
<tr>
<th>Default Source</th>
<th>Name</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diffusion</td>
<td>Brownian (<em>t^i) = \Pr \left( \exists s \in [t, t + \tau^D] : \begin{cases} \begin{array}{l} D</em>{s-}^i = 0 \land A_s^i = 1 \ \land (dN_s^i + \beta_j^i dN_s = 0) \end{array} \end{cases} \right)</td>
<td></td>
</tr>
<tr>
<td>Aggregate jump</td>
<td>agg_jump (<em>t^i) = \Pr \left( \exists s \in [t, t + \tau^D] : \begin{cases} \begin{array}{l} D</em>{s-}^i = 0 \land A_s^i \leq 1 \ \land \beta_j^i dN_s \neq 0 \end{array} \end{cases} \right)</td>
<td></td>
</tr>
<tr>
<td>Bank-specific jump</td>
<td>idio_jump (<em>t^i) = \Pr \left( \exists s \in [t, t + \tau^D] : \begin{cases} \begin{array}{l} D</em>{s-}^i = 0 \land A_s^i \leq 1 \ dN_s^i \neq 0 \end{array} \end{cases} \right)</td>
<td></td>
</tr>
<tr>
<td>Spillover</td>
<td>spillover (<em>t^i) = \Pr \left( \exists s \in [t, t + \tau^D] : \begin{cases} \begin{array}{l} D</em>{s-}^i = 0 \land A_s^i &gt; 1 \ \land A_{s+}^i \leq 1 \end{array} \end{cases} \right)</td>
<td></td>
</tr>
</tbody>
</table>

37
Table 2: Regression results in simulated data.

Panel (a) shows the parameter values and initial conditions used. 1,000 simulations were drawn, each with two years (504 trading days) worth of data. Draws resulting in a realized default are discarded. On each day, default probabilities are computed at the six month (126 trading day) horizon. Each default probability is computed as a Monte-Carlo average of 1,000 draws.

Panels (b) show the results of running regression equation (7) on of the generated data. Columns correspond to different left-hand-side variables, while rows correspond to right-hand-side variables. Non-zero values are evidence of a direct or indirect linkage. Recall:

\[
\Delta \text{jump}_{i+t+1}^{[i]} = \alpha[i] + \beta[i] \Delta \text{agg\_jump}_{t+1}^{[i]} + \chi[i] \Delta \text{Brownian}_{t+1}^{[i]} + \sum_{j \neq i} \gamma_{ij}^{[i]} \Delta \text{Brownian}_{t+1}^{[j]} + u_{t+1}^{[i]}
\]

(a) Parameter values and initial conditions.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Volatility</td>
<td>( \sigma )</td>
<td>Bank Volatility</td>
<td>( \sigma )</td>
</tr>
<tr>
<td>Aggregate Jump</td>
<td>( \kappa, \lambda, \sigma \lambda )</td>
<td>Bank Jump</td>
<td>( \kappa )</td>
</tr>
<tr>
<td>( \mu, \sigma \mu )</td>
<td>-0.20, 0.10</td>
<td>( \lambda )</td>
<td>(0.02, 0.02, 0.02)</td>
</tr>
<tr>
<td>( \sigma \lambda )</td>
<td>(0.10, 0.10, 0.10)</td>
<td>( \sigma _\lambda )</td>
<td>(0.10, 0.10, 0.10)</td>
</tr>
<tr>
<td>Bank Loadings ( \beta_i )</td>
<td>(1.00, 1.00, 1.00)</td>
<td>Initial Conditions</td>
<td>( \lambda_0 )</td>
</tr>
<tr>
<td>( \beta_j )</td>
<td>(1.00, 1.00, 1.00)</td>
<td>( \lambda_0 )</td>
<td>(1.15, 1.15, 1.15)</td>
</tr>
<tr>
<td>Spillovers ( \Theta )</td>
<td>( \begin{pmatrix} 0.20 &amp; 0.00 \ 0.00 &amp; 0.20 \end{pmatrix} )</td>
<td>( \lambda_0 )</td>
<td>(0.02, 0.02, 0.02)</td>
</tr>
</tbody>
</table>

(b) Regression (7).

<table>
<thead>
<tr>
<th>Jump</th>
<th>bank1</th>
<th>bank2</th>
<th>bank3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{Brownian} ) aggregate jump</td>
<td>0.19</td>
<td>0.22</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(0.77)</td>
<td>(0.62)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>bank1</td>
<td>-0.89</td>
<td>-0.02</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.07)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>bank2</td>
<td>0.24</td>
<td>0.24</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.29)</td>
<td>(0.31)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>bank3</td>
<td>0.70</td>
<td>0.77</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.33)</td>
<td>(0.31)</td>
<td>(0.01)</td>
</tr>
</tbody>
</table>
Table 3: Results for 2 years prior to March 13, 2008 (before collapse of Bear Stearns).

Columns correspond to the left-hand-side variable in regression equation (7). Rows correspond to the right-hand-side variables for each bank. Recall

\[
\Delta \text{jump}_t = \alpha_t + \beta_t \Delta \text{agg}_t + \chi_t \Delta \text{Brownian}_t + \sum_{j \neq i} \gamma_j \Delta \text{Brownian}_j + u_t
\]

Here, \(\text{jump}_t\) is the 6-month jump default risk extracted from CDS, options, and equity data, as outlined in Section 4. \(\text{Brownian}_t\) is the 6-month Brownian default risk extracted from the same data. \(\text{agg}_t\) is the 6-months probability of a jump estimated using S&P500 options. Point estimates are computed using ridge regression with a penalty parameter chosen optimally assuming non-zero coefficients are around 1 and that coefficients are non-zero with probability 0.3. The coefficients along the diagonal, as well as coefficients less than 0.1, are suppressed for readability. Row sums, in the last column, are computed using all coefficients except the diagonal.

The tickers are as follows: AIG (American International Group), BAC (Bank of America), BSC (Bear Stearns), C (Citigroup), COF (Capital One Financial), GS (Goldman Sachs), JPM (JPMorgan Chase), KEY (Keycorp), LEH (Lehman Brothers), MS (Morgan Stanley), WB (Wachovia), and WFC (Wells Fargo).

Warning: standard errors are the estimated sampling variation in ridge coefficients, and do not take into account the bias in the coefficients. Intervals formed by considering multiples of the standard error around the point estimates do not form valid confidence intervals. Standard errors are heteroskedasticity consistent.

<table>
<thead>
<tr>
<th>Brownian</th>
<th>AIG</th>
<th>BAC</th>
<th>BSC</th>
<th>C</th>
<th>COF</th>
<th>GS</th>
<th>JPM</th>
<th>KEY</th>
<th>LEH</th>
<th>MS</th>
<th>WB</th>
<th>WFC</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.01</td>
</tr>
<tr>
<td>BAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>BSC</td>
<td>1.48</td>
<td></td>
<td>0.48</td>
<td></td>
<td>0.46</td>
<td>0.25</td>
<td></td>
<td>0.24</td>
<td>0.42</td>
<td></td>
<td></td>
<td>3.37</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.52</td>
<td></td>
<td>1.23</td>
<td>0.12</td>
<td>0.14</td>
<td>0.19</td>
<td></td>
<td>-0.22</td>
<td>2.34</td>
</tr>
<tr>
<td>COF</td>
<td>0.52</td>
<td>-0.25</td>
<td>0.64</td>
<td></td>
<td></td>
<td>1.23</td>
<td>0.12</td>
<td>0.14</td>
<td>0.19</td>
<td></td>
<td></td>
<td>1.81</td>
<td></td>
</tr>
<tr>
<td>GS</td>
<td>0.08</td>
<td>-0.20</td>
<td>0.25</td>
<td></td>
<td>0.12</td>
<td>0.39</td>
<td>0.50</td>
<td>0.36</td>
<td>0.25</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JPM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KEY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.03</td>
<td></td>
</tr>
<tr>
<td>LEH</td>
<td>0.30</td>
<td>-0.45</td>
<td>0.29</td>
<td></td>
<td>1.44</td>
<td>0.29</td>
<td>0.18</td>
<td>0.29</td>
<td>0.64</td>
<td>0.72</td>
<td></td>
<td>3.26</td>
<td></td>
</tr>
<tr>
<td>MS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.30</td>
<td></td>
<td></td>
<td>0.18</td>
<td>0.29</td>
<td>0.64</td>
<td>0.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>WFC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.01</td>
<td></td>
</tr>
</tbody>
</table>

39
Table 4: Results for 2 years prior to March 17, 2008 (after collapse of Bear Stearns).

Columns correspond to the left-hand-side variable in regression equation (7). Rows correspond to the right-hand-side variables for each bank. Recall

\[
\Delta \text{jump}_t = \alpha_t + \beta \Delta \text{agg}_t + \chi \Delta \text{Brownian}_t + \sum_{j \neq i}^\gamma_j \Delta \text{Brownian}_j + \epsilon_t
\]

Here, \( \Delta \text{jump}_t \) is the 6-month jump default risk extracted from CDS, options, and equity data, as outlined in Section 4. \( \Delta \text{agg}_t \) is the 6-months probability of a jump estimated using S&P500 options. Point estimates are computed using ridge regression with a penalty parameter chosen optimally assuming non-zero coefficients are around 1 and that coefficients are non-zero with probability 0.3. The coefficients along the diagonal, as well as coefficients less than 0.1, are suppressed for readability. Row sums, in the last column, are computed using all coefficients except the diagonal.

The tickers are as follows: AIG (American International Group), BAC (Bank of America), BSC (Bear Stearns), C (Citigroup), COF (Capital One Financial), GS (Goldman Sachs), JPM (JPMorgan Chase), KEY (Keycorp), LEH (Lehman Brothers), MS (Morgan Stanley), WB (Wachovia), and WFC (Wells Fargo).

Warning: standard errors are the estimated sampling variation in ridge coefficients, and do not take into account the bias in the coefficients. Intervals formed by considering multiples of the standard error around the point estimates do not form valid confidence intervals. Standard errors are heteroskedasticity consistent.

<table>
<thead>
<tr>
<th></th>
<th>AIG</th>
<th>BAC</th>
<th>BSC</th>
<th>C</th>
<th>COF</th>
<th>GS</th>
<th>JPM</th>
<th>KEY</th>
<th>LEH</th>
<th>MS</th>
<th>WB</th>
<th>WFC</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Jump-Like</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Brownian</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIG</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAC</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BSC</td>
<td></td>
<td>0.13</td>
<td>0.36</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.73</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>COF</td>
<td>0.34</td>
<td>-0.21</td>
<td>1.00</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.49</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.19)</td>
<td>(1.22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS</td>
<td>0.40</td>
<td>0.98</td>
<td>-0.11</td>
<td>0.47</td>
<td>0.34</td>
<td>0.18</td>
<td>-0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.15</td>
</tr>
<tr>
<td></td>
<td>(0.33)</td>
<td>(0.66)</td>
<td>(0.39)</td>
<td>(0.28)</td>
<td>(0.31)</td>
<td>(0.28)</td>
<td>(0.39)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JPM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.08</td>
</tr>
<tr>
<td>KEY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.08</td>
</tr>
<tr>
<td>LEH</td>
<td>0.89</td>
<td>2.30</td>
<td>5.81</td>
<td>0.30</td>
<td>-0.16</td>
<td>2.04</td>
<td>0.38</td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11.96</td>
</tr>
<tr>
<td></td>
<td>(0.46)</td>
<td>(1.54)</td>
<td>(3.56)</td>
<td>(0.16)</td>
<td>(0.31)</td>
<td>(1.35)</td>
<td>(0.21)</td>
<td>(0.25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS</td>
<td>0.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.39</td>
<td></td>
<td></td>
<td>-0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.27)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WFC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>
Table 5: Results for 2 years prior to August 30, 2008.

Columns correspond to the left-hand-side variable in regression equation (7). Rows correspond to the right-hand-side variables for each bank. Recall

\[
\Delta \text{jump}_i^{[t+1]} = a_i^{[t]} + \beta_i^{[t]} \Delta \text{agg jump}_{t+1} + \chi_i^{[t]} \Delta \text{Brownian}_{t+1} + \sum_{j \neq i} \gamma_j^{[i]} \Delta \text{Brownian}_{t+1} + u_i^{[t+1]}
\]

Here, \( \Delta \text{jump}_i^{[t+1]} \) is the 6-month jump default risk extracted from CDS, options, and equity data, as outlined in Section 4. \( \Delta \text{Brownian}_{t+1} \) is the 6-month Brownian default risk extracted from the same data. \( \Delta \text{agg jump}_{t+1} \) is the 6-months probability of a jump estimated using S&P500 options. Point estimates are computed using ridge regression with a penalty parameter chosen optimally assuming non-zero coefficients are around 1 and that coefficients are non-zero with probability 0.3. The coefficients along the diagonal, as well as coefficients less than 0.1, are suppressed for readability. Row sums, in the last column, are computed using all coefficients except the diagonal.

The tickers are as follows: AIG (American International Group), BAC (Bank of America), C (Citigroup), COF (Capital One Financial), GS (Goldman Sachs), JPM (JPMorgan Chase), LEH (Lehman Brothers), MS (Morgan Stanley), WB (Wachovia), and WFC (Wells Fargo).

Warning: standard errors are the estimated sampling variation in ridge coefficients, and do not take into account the bias in the coefficients. Intervals formed by considering multiples of the standard error around the point estimates do not form valid confidence intervals. Standard errors are heteroskedasticity consistent.

<table>
<thead>
<tr>
<th>Brownian</th>
<th>AIG</th>
<th>BAC</th>
<th>C</th>
<th>COF</th>
<th>GS</th>
<th>JPM</th>
<th>LEH</th>
<th>MS</th>
<th>WB</th>
<th>WFC</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.04</td>
</tr>
<tr>
<td>BAC</td>
<td>0.24</td>
<td>0.72</td>
<td>0.26</td>
<td>0.27</td>
<td>0.35</td>
<td>-0.11</td>
<td>0.16</td>
<td></td>
<td></td>
<td></td>
<td>1.84</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.53)</td>
<td>(0.30)</td>
<td>(0.37)</td>
<td>(0.78)</td>
<td>(0.23)</td>
<td>(0.21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td>-0.20</td>
<td></td>
<td>-1.56</td>
<td>-0.17</td>
<td>0.14</td>
<td></td>
<td>-1.95</td>
<td></td>
<td></td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.12)</td>
<td></td>
<td>(1.04)</td>
<td>(0.16)</td>
<td>(0.61)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COF</td>
<td></td>
<td></td>
<td></td>
<td>0.37</td>
<td></td>
<td></td>
<td></td>
<td>0.33</td>
<td></td>
<td></td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.09)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS</td>
<td>0.79</td>
<td>0.37</td>
<td>1.19</td>
<td></td>
<td>2.09</td>
<td>0.39</td>
<td>2.13</td>
<td>0.43</td>
<td>1.04</td>
<td>8.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.55)</td>
<td>(0.39)</td>
<td>(0.87)</td>
<td></td>
<td>(1.34)</td>
<td>(0.88)</td>
<td>(1.64)</td>
<td>(0.24)</td>
<td>(0.86)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JPM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.56</td>
<td></td>
<td>0.24</td>
<td>-0.30</td>
<td></td>
<td></td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.36)</td>
<td>(0.50)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEH</td>
<td>0.27</td>
<td>0.28</td>
<td>1.08</td>
<td>0.32</td>
<td>0.17</td>
<td></td>
<td>0.11</td>
<td></td>
<td>0.10</td>
<td>2.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.26)</td>
<td>(0.60)</td>
<td>(0.22)</td>
<td>(0.17)</td>
<td></td>
<td>(0.06)</td>
<td>(0.09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS</td>
<td>-0.79</td>
<td>-0.19</td>
<td>0.64</td>
<td></td>
<td>-0.41</td>
<td>-0.17</td>
<td>-0.47</td>
<td>-1.47</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.36)</td>
<td>(0.13)</td>
<td>(0.51)</td>
<td></td>
<td>(0.68)</td>
<td>(0.33)</td>
<td>(0.34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.68</td>
<td>-0.16</td>
<td>0.13</td>
<td></td>
<td>0.14</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.49)</td>
<td>(0.21)</td>
<td>(0.20)</td>
<td></td>
<td>(0.16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WFC</td>
<td>0.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.08</td>
<td>-0.23</td>
<td></td>
<td>-1.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.68)</td>
<td>(0.66)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 6: Results for 2 years prior to September 12, 2008 (collapse of Lehman).

Columns correspond to the left-hand-side variable in regression equation (7). Rows correspond to the right-hand-side variables for each bank. Recall

\[
\Delta \text{jump}_i^{\text{like}}_{t+1} = \alpha_i + \beta_i \Delta \text{agg}_i \text{jump}_{t+1} + \chi_i \Delta \text{Brownian}_i^{\text{agg}}_{t+1} + \sum_{j \neq i} \gamma_j \Delta \text{Brownian}_j^{\text{agg}}_{t+1} + u_{t+1}
\]

Here, jump\_like\_i^{\text{like}}_{t+1} is the 6-month jump default risk extracted from CDS, options, and equity data, as outlined in Section 4. Brownian\_i^{\text{agg}}_{t+1} is the 6-month Brownian default risk extracted from the same data. \hat{\Delta} \text{jump}_{t+1} is the 6-months probability of a jump estimated using S&P500 options. Point estimates are computed using ridge regression with a penalty parameter chosen optimally assuming non-zero coefficients are around 1 and that coefficients are non-zero with probability 0.3. The coefficients along the diagonal, as well as coefficients less than 0.1, are suppressed for readability. Row sums, in the last column, are computed using all coefficients except the diagonal.

The tickers are as follows: AIG (American International Group), BAC (Bank of America), C (Citigroup), COF (Capital One Financial), GS (Goldman Sachs), JPM (JPMorgan Chase), LEH (Lehman Brothers), MS (Morgan Stanley), WB (Wachovia), and WFC (Wells Fargo).

Warning: standard errors are the estimated sampling variation in ridge coefficients, and do not take into account the bias in the coefficients. Intervals formed by considering multiples of the standard error around the point estimates do not form valid confidence intervals. Standard errors are heteroskedasticity consistent.

<table>
<thead>
<tr>
<th>Brownian</th>
<th>AIG</th>
<th>BAC</th>
<th>C</th>
<th>COF</th>
<th>GS</th>
<th>JPM</th>
<th>LEH</th>
<th>MS</th>
<th>WB</th>
<th>WFC</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIG</td>
<td>0.18</td>
<td>(0.17)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.25</td>
</tr>
<tr>
<td>BAC</td>
<td>-0.71</td>
<td>(0.37)</td>
<td>0.18</td>
<td>0.93</td>
<td>0.41</td>
<td>0.45</td>
<td>0.68</td>
<td>0.13</td>
<td></td>
<td>2.11</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>-0.23</td>
<td>(0.11)</td>
<td>-0.11</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>COF</td>
<td>-0.35</td>
<td>(0.13)</td>
<td></td>
<td>0.33</td>
<td>0.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>GS</td>
<td>0.65</td>
<td>(0.43)</td>
<td>0.73</td>
<td>1.06</td>
<td>0.23</td>
<td>2.25</td>
<td>0.76</td>
<td>2.30</td>
<td>0.30</td>
<td>1.07</td>
<td>9.35</td>
</tr>
<tr>
<td>JPM</td>
<td>-0.11</td>
<td>(0.12)</td>
<td>-0.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.16</td>
<td>0.12</td>
</tr>
<tr>
<td>LEH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>MS</td>
<td>-0.32</td>
<td>(0.45)</td>
<td>-0.32</td>
<td>-0.14</td>
<td>0.85</td>
<td>0.22</td>
<td>0.14</td>
<td>0.81</td>
<td>-0.15</td>
<td></td>
<td>1.05</td>
</tr>
<tr>
<td>WB</td>
<td>0.41</td>
<td>(0.42)</td>
<td></td>
<td>0.85</td>
<td>0.13</td>
<td>1.13</td>
<td></td>
<td>0.16</td>
<td></td>
<td>2.56</td>
<td></td>
</tr>
<tr>
<td>WFC</td>
<td>-0.13</td>
<td>(0.10)</td>
<td>0.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.20</td>
<td>-0.14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

42
Table 7: Results for 2 years prior to September 16, 2008 (AIG bailout).

Columns correspond to the left-hand-side variable in regression equation (7). Rows correspond to the right-hand-side variables for each bank. Recall

\[
\Delta \text{jump}_t = \alpha_t + \beta \Delta \text{agg jump}_t + \chi_t \Delta \text{Brownian}_t + \sum_{j \neq i} \gamma_j \Delta \text{Brownian}_j + u_t
\]

Here, \( \Delta \text{jump}_t \) is the 6-month jump default risk extracted from CDS, options, and equity data, as outlined in Section 4. \( \Delta \text{Brownian}_t \) is the 6-month Brownian default risk extracted from the same data. \( \Delta \text{agg jump}_t \) is the 6-months probability of a jump estimated using S&P500 options. Point estimates are computed using ridge regression with a penalty parameter chosen optimally assuming non-zero coefficients are around 1 and that coefficients are non-zero with probability 0.3. The coefficients along the diagonal, as well as coefficients less than 0.1, are suppressed for readability. Row sums, in the last column, are computed using all coefficients except the diagonal.

The tickers are as follows: AIG (American International Group), BAC (Bank of America), C (Citigroup), COF (Capital One Financial), GS (Goldman Sachs), JPM (JPMorgan Chase), MS (Morgan Stanley), WB (Wachovia), and WFC (Wells Fargo).

Warning: standard errors are the estimated sampling variation in ridge coefficients, and do not take into account the bias in the coefficients. Intervals formed by considering multiples of the standard error around the point estimates do not form valid confidence intervals. Standard errors are heteroskedasticity consistent.

<table>
<thead>
<tr>
<th>Jump-Like</th>
<th>AIG</th>
<th>BAC</th>
<th>C</th>
<th>COF</th>
<th>GS</th>
<th>JPM</th>
<th>MS</th>
<th>WB</th>
<th>WFC</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brownian</td>
<td>AIG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.13</td>
</tr>
<tr>
<td>AIG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAC</td>
<td>1.01</td>
<td>0.73</td>
<td>0.47</td>
<td>0.29</td>
<td>0.17</td>
<td>2.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(0.41)</td>
<td>(0.40)</td>
<td>(0.53)</td>
<td>(0.46)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.13</td>
<td>0.16</td>
<td>0.29</td>
<td>-0.71</td>
<td>0.13</td>
<td>-0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.14)</td>
<td>(0.39)</td>
<td>(0.69)</td>
<td>(0.16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COF</td>
<td>-0.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS</td>
<td>0.83</td>
<td>0.83</td>
<td>0.97</td>
<td>0.27</td>
<td>2.12</td>
<td>2.98</td>
<td>1.11</td>
<td>9.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.59)</td>
<td>(0.58)</td>
<td>(0.17)</td>
<td>(1.27)</td>
<td>(1.48)</td>
<td>(0.89)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JPM</td>
<td>1.56</td>
<td>0.28</td>
<td>-0.15</td>
<td>0.23</td>
<td>-0.11</td>
<td>-0.21</td>
<td>1.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.89)</td>
<td>(0.14)</td>
<td>(0.26)</td>
<td>(0.35)</td>
<td>(0.06)</td>
<td>(0.11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS</td>
<td>1.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>(0.75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.16)</td>
</tr>
<tr>
<td>WFC</td>
<td>0.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Columns correspond to the left-hand-side variable in regression equation (7). Rows correspond to the right-hand-side variables for each bank. Recall

\[
\Delta \text{jump\_like}_{i,t+1} = \alpha[i] + \beta[i] \Delta \text{agg\_jump}_{i,t+1} + \chi[i] \Delta \text{Brownian}_{i,t+1} + \sum_{j \neq i} \gamma[i] \Delta \text{Brownian}_{j,t+1} + u[i,t+1]
\]

Here, \(\text{jump\_like}_{i,t+1}\) is the 6-month jump default risk extracted from CDS, options, and equity data, as outlined in Section 4. \(\text{Brownian}_{i,t+1}\) is the 6-month Brownian default risk extracted from the same data. \(\Delta \text{agg\_jump}_{i,t+1}\) is the 6-months probability of a jump estimated using S&P500 options. Point estimates are computed using ridge regression with a penalty parameter chosen optimally assuming non-zero coefficients are around 1 and that coefficients are non-zero with probability 0.3. The coefficients along the diagonal, as well as coefficients less than 0.1, are suppressed for readability. Row sums, in the last column, are computed using all coefficients except the diagonal.

The tickers are as follows: AIG (American International Group), BAC (Bank of America), C (Citigroup), COF (Capital One Financial), GS (Goldman Sachs), JPM (JPMorgan Chase), MS (Morgan Stanley), WB (Wachovia), and WFC (Wells Fargo).

Warning: standard errors are the estimated sampling variation in ridge coefficients, and do not take into account the bias in the coefficients. Intervals formed by considering multiples of the standard error around the point estimates do not form valid confidence intervals. Standard errors are heteroskedasticity consistent.

<table>
<thead>
<tr>
<th>Brownian</th>
<th>AIG</th>
<th>BAC</th>
<th>C</th>
<th>COF</th>
<th>GS</th>
<th>JPM</th>
<th>MS</th>
<th>WB</th>
<th>WFC</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAC</td>
<td>0.72</td>
<td>0.92</td>
<td>0.39</td>
<td>0.10</td>
<td>0.17</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td>2.96</td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td>(0.74)</td>
<td>(0.37)</td>
<td>(0.30)</td>
<td>(0.12)</td>
<td>(0.38)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.64</td>
<td></td>
<td></td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.36)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COF</td>
<td>-0.93</td>
<td>0.53</td>
<td>0.86</td>
<td>0.20</td>
<td>0.91</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td>1.87</td>
</tr>
<tr>
<td></td>
<td>(0.67)</td>
<td>(0.42)</td>
<td>(0.36)</td>
<td>(0.15)</td>
<td>(0.62)</td>
<td>(0.11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS</td>
<td>-0.30</td>
<td>-0.14</td>
<td>0.25</td>
<td>-0.14</td>
<td>-0.18</td>
<td>0.65</td>
<td>0.25</td>
<td></td>
<td></td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.13)</td>
<td>(0.37)</td>
<td>(0.13)</td>
<td>(0.22)</td>
<td>(0.49)</td>
<td>(0.22)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JPM</td>
<td>0.17</td>
<td>1.31</td>
<td>0.40</td>
<td>0.67</td>
<td>2.04</td>
<td>0.14</td>
<td>0.25</td>
<td></td>
<td></td>
<td>5.04</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(1.25)</td>
<td>(0.18)</td>
<td>(0.39)</td>
<td>(0.76)</td>
<td>(0.33)</td>
<td>(0.19)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.23</td>
<td></td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WB</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.49</td>
<td></td>
<td></td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.47)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WFC</td>
<td>0.43</td>
<td>-0.18</td>
<td>0.90</td>
<td>-0.19</td>
<td>-0.36</td>
<td>-0.32</td>
<td>1.23</td>
<td>0.13</td>
<td></td>
<td>1.64</td>
</tr>
<tr>
<td></td>
<td>(0.51)</td>
<td>(0.31)</td>
<td>(0.63)</td>
<td>(0.26)</td>
<td>(0.45)</td>
<td>(0.34)</td>
<td>(0.83)</td>
<td>(0.27)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 9: Results for 2 years before December 31, 2017.

Columns correspond to the left-hand-side variable in regression equation (7). Rows correspond to the right-hand-side variables for each bank. Recall

(7) $\Delta \text{jump}_i[t+1] = \alpha_i[t] + \beta_i[t] \Delta \text{agg}_i[t+1] + \chi_i[t] \Delta \text{Brownian}_i[t+1] + \sum_{j \neq i} \gamma_j[t] \Delta \text{Brownian}_j[t+1] + u_i[t+1]$

Here, $\text{jump}_i[t+1]$ is the 6-month jump default risk extracted from CDS, options, and equity data, as outlined in Section 4. $\text{Brownian}_i[t+1]$ is the 6-month Brownian default risk extracted from the same data. $\hat{\text{agg}}_i[t+1]$ is the 6-months probability of a jump estimated using S&P500 options. Point estimates are computed using ridge regression with a penalty parameter chosen optimally assuming non-zero coefficients are around 1 and that coefficients are non-zero with probability 0.3. Row sums, in the last column, are computed using all coefficients except the diagonal.

The tickers are as follows: AXP (American Express), BAC (Bank of America), C (Citigroup), COF (Capital One Financial), GS (Goldman Sachs), JPM (JPMorgan Chase), KEY (KeyCorp), MS (Morgan Stanley), SCHW (Charles Schwab), USB (US Bancorp), and WFC (Wells Fargo).

Warning: standard errors are the estimated sampling variation in ridge coefficients, and do not take into account the bias in the coefficients. Intervals formed by considering multiples of the standard error around the point estimates do not form valid confidence intervals. Standard errors are heteroskedasticity consistent.

<table>
<thead>
<tr>
<th>Brownian</th>
<th>AXP</th>
<th>BAC</th>
<th>C</th>
<th>COF</th>
<th>GS</th>
<th>JPM</th>
<th>KEY</th>
<th>MS</th>
<th>SCHW</th>
<th>USB</th>
<th>WFC</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agg. Jump</td>
<td>0.06</td>
<td>0.25</td>
<td>0.23</td>
<td>0.04</td>
<td>0.21</td>
<td>0.21</td>
<td>0.01</td>
<td>0.33</td>
<td>-0.01</td>
<td>0.05</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.08)</td>
<td>(0.09)</td>
<td>(0.03)</td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.01)</td>
<td>(0.07)</td>
<td>(0.01)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td>AXP</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>BAC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>C</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>COF</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>GS</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>JPM</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>KEY</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>MS</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>SCHW</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>USB</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>WFC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
</tbody>
</table>
Appendices

A  Data and Estimation

B  Spillover Equilibria

Equation (3) potentially has multiple solutions, corresponding to the idea of self-fulfilling defaults. In this section, I characterize the range of solutions, and show an algorithm to compute the most conservative one. The arguments are heavily reminiscent of Eisenberg and Noe (2001).

For notation, it is convenient to stack the equations to get one vector-valued equation. Let $A_t = (A^1_t \cdots A^N_t)'$. Then, equation (3) can be rewritten as

(B.1) $A_{t+} = A_t - \Theta \{A_{t+} \leq 1\}$

where $\Theta = (\theta_{i,j})_{i,j}$.

Before proceeding, it is helpful at this point to consider an example. Suppose $N = 2$, and suppose $\theta_{1,2} > 0$ and $\theta_{2,1} > 0$. Then, for $A_t$ strictly greater than but sufficiently close to 1, there are two solutions to equation (B.1). The first is $A_{t+} = A_t$. This corresponds to the equilibrium I focus on in this paper: neither bank defaults on its own, so no spillovers occur. The second is $A_{t+} = A_t - \Theta 1$. Although neither starts below its default boundary, both are sufficiently close to their default boundary that they will default if the other bank defaults. This equilibrium is effectively a self-fulfilling default: bank 1 defaults because bank 2 defaults and vice-versa.

The first proposition guarantees the existence of a solution to (B.1). Furthermore, it tells us that there are “best” and “worst” equilibria that bound all the others.

Proposition B.1. Equation (B.1) has at least one solution. Furthermore, there exist two not necessarily distinct solutions to (B.1), labeled $A^+_{t+}$ and $A^-_{t+}$, such that any solution $A_{t+}$ to (B.1) satisfies $A^-_{t+} \leq A_{t+} \leq A^+_{t+}$.

Proof. Let $L$ be given by the following product of closed intervals

$L = \prod_{i=1}^{N} [A^i_t - \sum_{j \neq i} \theta_{i,j}, A^i_t]$
Let $T : L \to L$ be defined by

$$T(x) = A_t - \Theta 1 \{ x \leq 1 \}$$

I claim that $T$ is order-preserving. To see this, note that if $x, y \in L$ with $x \leq y$, then $1 \{ x \leq 1 \} \geq 1 \{ y \leq 1 \}$. Nonnegativity of $\Theta$ implies $-\Theta 1 \{ x \leq 1 \} \leq -\Theta 1 \{ y \leq 1 \}$, from which it follows that $T(x) \leq T(y)$.

The Knaster-Tarski theorem (Tarski’s fixed point theorem) gives us that the set of fixed points of $T$ is a complete lattice. Hence, it is non-empty, giving us the existence of a solution. Furthermore, it means that the fixed points of $T$ have a supremum and and infemum, and that these are themselves fixed points. These two extrema are $A_{t+}^+$ and $A_{t+}^-$, respectively.

An immediate corollary to this proposition is that these two equilibria also bound which banks default for all equilibria. If a bank defaults in the “+” equilibrium, it defaults in all equilibria. Similarly, if a bank survives in the “−” equilibrium, it survives in all equilibria. Put differently, the “+” equilibrium has minimal defaults, and the “−” equilibrium has maximal defaults, both in terms of the total number of defaults and the identities of the defaults.

**Corollary B.2.** If $A_{t+}^{+[i]} \leq 1$, then bank $i$ defaults in all equilibria. If $A_{t+}^{-[i]} > 1$, then bank $i$ does not default in any equilibrium.

The next proposition tells us how to compute $A_{t+}^\pm$. It tells us that we can just start with $A_t$ and “push it through” equation (B.1) at most $N$ times to obtain the fixed point.

**Proposition B.3.** Let $A_t^{(0)} = A_t$, and define $A_t^{(n+1)} = A_t - \Theta 1 \{ A_t^{(n)} \leq 1 \}$ recursively for $n \geq 0$. Then, $A_{t+}^+ = A_t^{(N)}$.

**Proof.** Let $L$ and $T$ be defined as in the proof for Proposition B.1. Note that $A_t^{(n+1)} = T(A_t^{(n)})$.

First, I claim that $A_{t+}^+ \leq A_t^{(n)}$ for all $n \geq 0$. For $n = 0$, this is obvious:

$$A_{t+}^+ = T(A_{t+}^+) = A_t - \Theta 1 \{ A_{t+}^+ \leq 1 \} \leq A_t = A_t^{(0)}$$

Now, suppose it were true for some $n \geq 0$. Then $A_{t+}^+ \leq A_t^{(n)}$. Since $T$ is order-preserving, $T(A_{t+}^+) \leq T(A_t^{(n)})$. Using $T(A_t^{(n)}) = A_t^{(n+1)}$ and that $A_{t+}^+$ is a fixed point of $T$, we get that $A_{t+}^+ \leq A_t^{(n+1)}$. Therefore, by induction $A_{t+}^+ \leq A_t^{(n)}$ for all $n \geq 0$.

Second, I claim that $A_t^{(N)}$ is a fixed point of $T$. To see this, define $D^{(n)} = 1 \{ A_t^{(n)} \leq 1 \}$. Then, $A_t^{(n+1)} = A_t - \Theta D^{(n)}$. The sequence $A_t^{(n)}$ is non-increasing, from which it follows that the sequence $D^{(n)}$ is non-decreasing. Any non-decreasing sequence in $\{0, 1\}^N$ must have a repeated element after at most $N + 1$ iterations. Hence, for some $j \leq N$ we have that $D^{(j)} = D^{(j+1)}$. In fact, this must be true for some $j \leq N - 1$. The only way for all the $D^{(0)}, \ldots, D^{(N+1)}$ to be distinct is if $D^{(0)} = 0$, in which case $A_t^{(1)} = A_t^-$ and $D^{(1)} = D^{(0)}$.}

47
This in turn implies that $A_t^{(j+1)} = A_t^{(j+2)}$, which implies $A_t^{(j+1)}$ is a fixed point of $T$. This implies $A_t^{(n)} = A_t^{(j+1)}$ for all $n \geq j + 1$. Since $j \leq N - 1$, this implies $A_t^{(N)}$ is a fixed point of $T$. Proposition B.1 therefore gives us that $A_t^{(N)} \leq A_{t+}^+$.

Therefore, combining the two statements we have that $A_{t+}^+ \leq A_t^{(N)} \leq A_{t+}^+$. □

Although not particularly useful, $A_{t+}^−$ can be computed in an analogous way. We can start with $A_t^{(0)} = A_t - \Theta 1$ and push it through equation (B.1) at most $N$ times.

Though not used in this paper, it is also worth noting the linkage between the existence of multiple solutions to equation (B.1) and cycles in the directed graph with edges encoded by $\Theta$.

### B.1 Allowing for Self-Fulfilling Defaults

Equation (3) generally allows for multiple solutions, corresponding to multiple equilibria. In this paper, I have chosen the most conservative equilibrium, i.e., the equilibrium featuring the minimal amount of spillovers. The first spillovers are triggered by one or more banks falling below their default boundaries for reasons other than spillovers. These may cause further spillovers, but only if the initial spillovers caused additional banks to fall below their default boundaries.

Although it substantially complicates the theoretical analysis, the model behaves similarly with self-fulfilling default equilibria, as long as these equilibria are well behaved. Specifically, the equilibrium should be time invariant, and should be monotonic. To explain what this means, I first need to introduce some new notation.

Define an equilibrium to be a function $f$ that maps $A_t$ to $A_{t+}$ in a manner consistent with equation (B.1). Specifically, it is a function $f : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ such that

$$f(t, A) = A - \Theta 1 \{ f(t, A) \leq 1 \}$$

for all $t \in \mathbb{R}$ and $A \in \mathbb{R}^N$. Note that I am only allowing for deterministic equilibria; it is not a random variable.

I call an equilibrium $f$ time invariant over time if it does not depend on time, i.e., if $f(t, A) = f(s, A)$ for all $s, t$, and $A$. I call an equilibrium monotonic if it is weakly increasing in $A$, i.e., $A \leq A'$ implies $f(t, A) \leq f(t, A')$ for all $t$, $A$, and $A'$. As an example, the function that maps $A_t$ to $A_{t+}^+$ is a monotonic and time invariant equilibrium.

Given an equilibrium $f$, for each bank $i$ and time $t$ we can define a default set $D_{f_i}^{[i]}$ as follows

$$D_{f_i}^{[i]} = \{ A \in \mathbb{R}^N | f(t, A) \cdot e_i \leq 1 \}$$

In words, this is the set asset levels (for all banks) for which bank $i$ defaults at time $t$. If the equilibrium is time invariant, the default set is also time invariant. If the equilibrium $f$ is monotonic,
then $D[f,i]$ is a lower set, i.e., it has the property that for any $A \in D[f,i]$ and for any $A' \leq A$, $A' \in D[f,i]$. There is effectively a “boundary” that when crossed causes bank $i$ to default.

We can characterize this set even more by setting the level of assets at all the other banks. Given $A[-i] \in \mathbb{R}^{N-1}$, and given a monotonic equilibrium $f$, the properties above guarantee there exists a cutoff level of assets at bank $i$ where the bank defaults if assets are below and survives if assets are above. I denote this level by

$$D[f,i](A[-i]) = \sup_{A[i]} \{ A = (A[1], \ldots, A[i], \ldots, A[N]) \mid A \in D[f,i] \}$$

Given a monotonic $f$, this function is weakly decreasing in $A[-i]$.

## C Linearizing the Formal Model

In this appendix section, I linearly approximate the formal model presented in Section 2.1-2.2 in the case of two banks. I start by considering a straightforward informal approximation that captures much of the intuition without the need for derivations. I then simplify the formal model, and linearize the resulting approximation. I compute the linear approximation to the spillover default risk at bank 1, and show that when $\theta_{1,2} > 0$, in a broad set of cases the coefficients on Brownian and idiosyncratic jump risk at bank 2 are positive.

### C.1 Informal Approximation

In general, we can approximate the full model in equation (2) by focusing on the case where there are two banks and at most one jump in the $\tau^D$ units of time. Let $p[1]$ denote the probability of an idiosyncratic jump at bank 1, $p[2]$ denote the probability of an idiosyncratic jump at bank 2, $p^{agg}$ denote the probability of an aggregate jump, and $p^{none}$ denote the remaining probability (corresponding to no jump). Similarly, let Pr$[1]$, Pr$[2]$, Pr$^{agg}$ and Pr$^{none}$ denote probabilities conditioning on the corresponding events. Let $A_{\tau^D}$ denote the level of assets relative to liabilities at time $\tau^D$ after any jumps and Brownian shocks but before computing spillovers. Then

$$\text{spillover}^{[1]} = p[1] \Pr[1][1 < A_{\tau^D} \leq 1 + \theta_{1,2}, A_{\tau^D}^{[2]} \leq 1]$$
$$+ p[2] \Pr[2][1 < A_{\tau^D} \leq 1 + \theta_{1,2}, A_{\tau^D}^{[2]} \leq 1]$$
$$+ p^{agg} \Pr^{agg}[1 < A_{\tau^D} \leq 1 + \theta_{1,2}, A_{\tau^D}^{[2]} \leq 1]$$
$$+ p^{none} \Pr^{none}[1 < A_{\tau^D} \leq 1 + \theta_{1,2}, A_{\tau^D}^{[2]} \leq 1]$$
For each probability, we can use Bayes’ law to write \( \Pr[x_1 < A_{\tau_D} \leq 1 + \theta_{1,2}, A_{\tau_D} \leq 1] = \Pr[x_1 \leq 1 + \theta_{1,2}, A_{\tau_D} \leq 1] \Pr[A_{\tau_D} \leq 1] \). If we combine this with

\[
\text{idio\_jump}^2 = p^2 \Pr[A_{\tau_D} \leq 1] \\
\text{agg\_jump}^2 = p^{agg} \Pr[A_{\tau_D} \leq 1] \\
\text{Brownian}^2 = p^{[1]} \Pr[A_{\tau_D} \leq 1] + p^{\text{none}} \Pr[A_{\tau_D} \leq 1]
\]

we obtain

\[
\text{spillover}^1 = \Pr[1 < A_{\tau_D} \leq 1 + \theta_{1,2}, A_{\tau_D} \leq 1] \times \text{idio\_jump}^2 \\
+ \Pr^{agg}[1 < A_{\tau_D} \leq 1 + \theta_{1,2}, A_{\tau_D} \leq 1] \times \text{agg\_jump}^2 \\
+ \Pr^{[1]}[1 < A_{\tau_D} \leq 1 + \theta_{1,2}, A_{\tau_D} \leq 2] p^{[1]} \Pr[A_{\tau_D} \leq 1] \\
+ \Pr^{\text{none}}[1 < A_{\tau_D} \leq 1 + \theta_{1,2}, A_{\tau_D} \leq 2] p^{\text{none}} \Pr[A_{\tau_D} \leq 1]
\]

Note that \( \Pr^{[1]} \) is the probability conditional on a jump occurring at bank 1, while \( \Pr^{\text{none}} \) is the probability conditional on no jump. Unless the jump intensity at bank 1 correlates with diffusive shocks at bank 2, we have

\[
\Pr^{[1]}[A_{\tau_D} \leq 1] = \Pr^{\text{none}}[A_{\tau_D} \leq 1]
\]

Under this assumption,

\[
\text{Brownian}^2 = (p^{[1]} + p^{\text{none}}) \Pr^{\text{none}}[A_{\tau_D} \leq 1]
\]

in which case we can write

\[
\text{spillover}^1 = \Pr[1 < A_{\tau_D} \leq 1 + \theta_{1,2}, A_{\tau_D} \leq 1] \times \text{idio\_jump}^2 \\
+ \Pr^{agg}[1 < A_{\tau_D} \leq 1 + \theta_{1,2}, A_{\tau_D} \leq 1] \times \text{agg\_jump}^2 \\
+ \left( \Pr^{[1]}[1 < A_{\tau_D} \leq 1 + \theta_{1,2}, A_{\tau_D} \leq 2] \frac{p^{[1]}}{p^{[1]} + p^{\text{none}}} \right) \\
+ \left( \Pr^{\text{none}}[1 < A_{\tau_D} \leq 1 + \theta_{1,2}, A_{\tau_D} \leq 2] \frac{p^{\text{none}}}{p^{[1]} + p^{\text{none}}} \right) \times \text{Brownian}^2
\]

This appears linear to be linear in each of the variables \( \text{idio\_jump}^2 \), \( \text{agg\_jump}^2 \), and \( \text{Brownian}^2 \). The coefficients of this “linearization” are conditional probabilities: conditional on the respective type of default happening at bank 2, what is the probability that bank 1 has assets between 1 and \( 1 + \theta_{1,2} \). Only in this region do spillovers lead to a default.

Obviously, the conditional probabilities in equation (C.1) are not actually constants, and hence it is not a true linearization. They depend on the all of the state variables of the system. State variables from bank 1 can easily be handled, in the linear sense, by adding other default probabilities from bank 1. State variables from bank 2 could, in principle, be more problematic. If changing
conditions at bank 2 substantially affect the conditional probabilities, in a linearized setup the
coefficients on bank 2’s default probabilities could have negative signs.

However, this type of sign flip is unlikely to happen. State variables at bank 2 only affects
the the conditional probabilities through the conditioning information. To get large changes, this
requires that information about bank 2 substantially affects the likelihood of bank 1 being in its
sensitive region (assets between 1 and 1 + \( \theta_{1,2} \)) conditional on bank 2 defaulting. This is only
possible if the two banks are highly correlated, and conditional on bank 2 defaulting bank 1 is very
close to

\[
C.2 \text{ Approximation to the Formal Model}
\]

In order to linearize the formal model, we need to compute derivatives of default probabilities with
respect to different state variables. The general model, even for very simple stochastic processes,
does not allow for these derivatives to be readily computed. I therefore simplify the model in
three ways. First, I make volatilities and jump-intensities constant. Together with a log-normal
specification for assets, this makes the distributions of assets log-normal. Second, I compute all
defaults at the time of maturity \( \tau_D \), including any spillovers. Third, I allow for at most one type
of jump to happen in the maturity time horizon \( \tau_D \).

The simpler version of the model starts by considering fixed volatilities and jump intensities in
(2). This gives us that assets evolve according to

\[
\frac{dA_t^{[i]}}{A_t^{[i]}} = \beta_{[i]} \sigma dZ_t + \beta_{[i]} \left( d\left( \sum_{j=1}^{N(t)} (V_j - 1) \right) + \mathbb{E}[1 - V] \lambda dt \right) \\
+ \sigma^{[i]} dZ_t^{[i]} + d\left( \sum_{j=1}^{N(t)} (V_j^{[i]} - 1) \right) + \mathbb{E}[1 - V^{[i]}] \lambda d t
\]

where \( N(t) \) is a Poisson process with constant arrival rate \( \lambda \) and \( N(t)^{[i]} \) is a Poisson process with
constant arrival rate \( \lambda^{[i]} \). Putting this in log terms

\[
d \log A_t^{[i]} = \left( -\frac{1}{2} \beta^2 (\sigma^{[i]})^2 - \frac{1}{2} \sigma^2 + \beta_{[i]} \mathbb{E}[1 - V] \lambda + \mathbb{E}[1 - V^{[i]}] \lambda^{[i]} \right) dt \\
+ d \left( \sum_{j=1}^{N(t)} \log (1 - \beta_{[j]}^{[i]} + \beta_{[j]}^{[i]} V_j) \right) + d \left( \sum_{j=1}^{N(t)} \log V_j^{[i]} \right)
\]

As with the empirical specification, let \( V_j \sim \log N(\mu_j, \sigma_j^2) \), and \( V_j^{[i]} \sim N(\mu_j^{[i]}, (\sigma_j^{[i]})^2) \). As a simplification, I approximate

\[
\log \left( (1 - \beta_{[j]}^{[i]} + \beta_{[j]}^{[i]} V_j) \right) \text{ app.} \approx N(\beta_{[j]}^{[i]} \mu_j, (\beta_{[j]}^{[i]} \sigma_j)^2)
\]
to ensure that assets have a log-normal distribution conditional on the number (and type) of jumps that have occurred.

I then allow for at most one jump to occur during the $\tau^D$ units of time. This is effectively an approximation around small jump intensities, in which case the probability of multiple jumps (either of the same or different types) is negligible. In this case, we can define the following useful jump-related quantities

\[
p^{[1]} \equiv \Pr[\text{idio. jump at } 1] = \frac{\lambda^{[1]}}{\lambda^{[1]} + \lambda^{[2]} + \lambda} \left[ 1 - \exp \left\{ - (\lambda^{[1]} + \lambda^{[2]} + \lambda) \tau^D \right\} \right]
\]

\[
p^{[2]} \equiv \Pr[\text{idio. jump at } 2] = \frac{\lambda^{[2]}}{\lambda^{[1]} + \lambda^{[2]} + \lambda} \left[ 1 - \exp \left\{ - (\lambda^{[1]} + \lambda^{[2]} + \lambda) \tau^D \right\} \right]
\]

\[
p^{\text{agg}} \equiv \Pr[\text{agg. jump}] = \frac{\lambda_{\text{agg}}}{\lambda^{[1]} + \lambda^{[2]} + \lambda} \left[ 1 - \exp \left\{ - (\lambda^{[1]} + \lambda^{[2]} + \lambda) \tau^D \right\} \right]
\]

\[
p^{\text{none}} \equiv \Pr[\text{no jump}] = \exp \left\{ - (\lambda^{[1]} + \lambda^{[2]} + \lambda) \tau^D \right\}
\]

For notational simplicity, also define the following useful time trend and diffusion related quantities

\[
\mu^{[i]} = -\frac{1}{2} (\beta_B^{[i]} \sigma^{[i]})^2 - \frac{1}{2} \sigma^2 + \beta_B^{[i]} \mathbb{E}[1 - V] \lambda + \mathbb{E}[1 - V^{[i]}] \lambda^{[i]}
\]

\[
\hat{\sigma}^{[i]} = \sqrt{(\sigma^{[i]})^2 + (\beta_B^{[i]} \sigma)^2}
\]

and, define the following two means

\[
\xi^{[1]} = \log A^{[1]}_0 + \mu^{[1]} \tau^D
\]

\[
\xi^{[2]} = \log A^{[2]}_0 + \mu^{[2]} \tau^D
\]

The quantity $\xi^{[i]}$ is expected level of assets at time $\tau^D$ prior to any spillovers and conditional on no jump occurring, and can be interpreted as a measure of the asset buffer that bank $i$ has.

Conditional on event $x$ occurring, where here $x \in \{[1], [2], \text{agg}, \text{none}\}$ denotes an idiosyncratic jump at bank 1, an idiosyncratic jump at bank 2, an aggregate jump, or no jump, respectively, the distribution of log assets at the two banks before computing any spillovers is given by $(\log A^{[1]}_{\tau D}, \log A^{[2]}_{\tau D}) | x \sim N(\nu^x, \Sigma^x)$, where

\[
\nu^{[1]} = \begin{pmatrix} \xi^{[1]} + \mu^{[1]}_J \\ \xi^{[2]} + \mu^{[2]}_J \end{pmatrix} \qquad \Sigma^{[1]} = \begin{pmatrix} \sigma^{[1]}_2 + \beta^{[1]} B \beta^{[2]} B \sigma^2 \tau^D \\ \beta^{[1]} B \beta^{[2]} B \sigma^2 \tau^D \end{pmatrix}
\]

\[
\nu^{[2]} = \begin{pmatrix} \xi^{[1]} \\ \xi^{[2]} + \mu^{[2]}_J \end{pmatrix} \qquad \Sigma^{[2]} = \begin{pmatrix} \sigma^{[1]}_2 + \beta^{[1]} B \beta^{[2]} B \sigma^2 \tau^D \\ \beta^{[1]} B \beta^{[2]} B \sigma^2 \tau^D \end{pmatrix}
\]

\[
\nu^{\text{agg}} = \begin{pmatrix} \xi^{[1]} + \beta^{[1]}_J \mu_J \\ \xi^{[2]} + \beta^{[2]}_J \mu_J \end{pmatrix} \qquad \Sigma^{\text{agg}} = \begin{pmatrix} \sigma^{[1]}_2 + \beta^{[1]} B \sigma^2 \tau^D + \beta^{[1]} B \beta^{[2]} B \sigma^2 \tau^D + \beta^{[1]}_J \beta^{[2]}_J \sigma^2 \\ \beta^{[1]} B \beta^{[2]} B \sigma^2 \tau^D + \beta^{[1]}_J \beta^{[2]}_J \sigma^2 \end{pmatrix}
\]

\[
\nu^{\text{none}} = \begin{pmatrix} \xi^{[1]} \\ \xi^{[2]} \end{pmatrix} \qquad \Sigma^{\text{none}} = \begin{pmatrix} \sigma^{[1]}_2 + \beta^{[1]} B \sigma^2 \tau^D \\ \beta^{[1]} B \beta^{[2]} B \sigma^2 \tau^D \end{pmatrix}
\]
As am I only considering the values of $A^{[i]}$ at time $\tau^D$, the two jump probabilities and the Brownian probability are straightforward

\[
\text{idio\_jump}^{[i]} = p^{[i]} \Phi \left( -\frac{\nu_i^{[i]}}{\sqrt{\Sigma_{ii}}} \right)
\]
\[
\text{agg\_jump}^{[i]} = p_{agg}^{[i]} \Phi \left( -\frac{\nu_{agg}^{[i]}}{\sqrt{\Sigma_{aggii}}} \right)
\]
\[
\text{Brownian}^{[i]} = p^{[-i]} \Phi \left( -\frac{\nu_i^{[-i]}}{\sqrt{\Sigma_{ii}}} \right) + p^{\text{none}} \Phi \left( -\frac{\nu_i^{\text{none}}}{\sqrt{\Sigma_{ii}}} \right)
\]

(C.2)

To get an expression for the spillover default probability, let $F(x_1, x_2; \mu, \Sigma)$ denote the CDF of the bivariate normal $N(\mu, \Sigma)$. Then

(C.3) $\text{spillover}^{[1]}_t = \sum_{x \in \{1, 2, \text{agg}, \text{none}\}} p^x (F(\log(1 + \theta_{1,2}), 0; \nu^x, \Sigma^x) - F(0, 0; \nu^x, \Sigma^x))$

C.3 A Note on the Bivariate Normal

Before linearizing the system, it is helpful to know the derivatives of the CDF or a bivariate normal distribution. As before let $F(\cdot, \cdot; \mu, \Sigma)$ denote the CDF of a bivariate normal distribution with mean $\mu$ and variance matrix $\Sigma$. Let $\phi(\cdot, \cdot; \mu, \Sigma)$ denote the corresponding density function, and let $\phi(\cdot; \mu, \sigma^2)$ denote the density function of a univariate normal distribution with mean $\mu$ and variance $\sigma^2$. Then

\[
\frac{\partial F(x_1, x_2; \mu, \Sigma)}{\partial x_1} = \int_{-\infty}^{x_2} \phi(x_1, t_2; \mu, \Sigma)dt_2
\]
\[
= \phi(x_1; \mu_1, \Sigma_{11}) \Phi \left( \frac{x_2 - (\mu_2 + \frac{\Sigma_{12}}{\Sigma_{11}}(x_1 - \mu_1))}{\sqrt{\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}}} \right)
\]
\[
\frac{\partial F(x_1, x_2; \mu, \Sigma)}{\partial x_2} = \phi(x_2; \mu_2, \Sigma_{22}) \Phi \left( \frac{x_1 - (\mu_1 + \frac{\Sigma_{12}}{\Sigma_{22}}(x_2 - \mu_2))}{\sqrt{\Sigma_{11} - \frac{\Sigma_{12}^2}{\Sigma_{22}}}} \right)
\]
In particular, this means that

\[ F_1(x_1', x_2; \mu, \Sigma) - F_1(x_1, x_2; \mu, \Sigma) = \phi(x_1'; \mu_1, \Sigma_{11}) \Phi \left( \frac{x_2 - \left( \mu_2 + \frac{\Sigma_{12}}{\Sigma_{11}} (x_1' - \mu_1) \right)}{\sqrt{\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}} \right) - \phi(x_1; \mu_1, \Sigma_{11}) \Phi \left( \frac{x_2 - \left( \mu_2 + \frac{\Sigma_{12}}{\Sigma_{11}} (x_1 - \mu_1) \right)}{\sqrt{\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}} \right) \]

\[ F_2(x_1', x_2; \mu, \Sigma) - F_2(x_1, x_2; \mu, \Sigma) = \phi(x_2; \mu_2, \Sigma_{22}) \left[ \Phi \left( \frac{x_1' - \left( \mu_1 + \frac{\Sigma_{12}}{\Sigma_{22}} (x_2 - \mu_2) \right)}{\sqrt{\Sigma_{11} - \frac{\Sigma_{12}^2}{\Sigma_{22}}} \right) - \Phi \left( \frac{x_1 - \left( \mu_1 + \frac{\Sigma_{12}}{\Sigma_{22}} (x_2 - \mu_2) \right)}{\sqrt{\Sigma_{11} - \frac{\Sigma_{12}^2}{\Sigma_{22}}} \right) \right] \]

Suppose \( x_1' > x_1 \). The first expression, the difference in partials w.r.t. \( x_1 \), has an ambiguous sign. If \( \Sigma_{12} > 0 \), the first \( \Phi(\cdot) \) term is smaller than the second \( \Phi(\cdot) \) term, tending to push the quantity negative. However, the different weights given by the \( \phi(\cdot) \) terms can flip this sign. The second expression, the difference in partials w.r.t. \( x_2 \), has an unambiguous sign. It is always positive.

### C.4 Linearizing the Approximate Model

Now we can linearize the whole system of equations (C.2) and (C.3). One way to do this is to work in the state variables of the model: \( \log A_0^{[1]} \), \( \log A_0^{[2]} \), \( \lambda^{[1]} \), \( \lambda^{[2]} \), and \( \lambda \). However, this is quite ugly, and ultimately these are not necessary. Instead, I will linearize in the transformed state variables \( \xi^{[1]} \), \( \xi^{[2]} \), \( p^{[1]} \), \( p^{[2]} \), \( p^{agg} \), as the system is already substantially more linear in these variables. Note that there is a bijection between these two sets of variables, and the Jacobian of this transformation does not vanish, so linearizing in each set yields the same results.

I will start with the easy equations (C.2). Omitted derivatives are zero. The derivatives are
given by:

\[
\frac{\partial \text{idio}_\text{jump}^{[i]}}{\partial p^{[i]}} = \Phi \left( -\frac{\nu_i^{[i]}}{\sqrt{\Sigma[i]_{ii}}} \right) > 0
\]

\[
\frac{\partial \text{idio}_\text{jump}^{[i]}}{\partial \xi^{[i]}} = -p^{[i]} \phi \left( -\frac{\nu_i^{[i]}}{\sqrt{\Sigma[i]_{ii}}} \right) \frac{1}{\sqrt{\Sigma[i]_{ii}}} < 0
\]

\[
\frac{\partial \text{agg}_\text{jump}^{[i]}}{\partial p^{\text{agg}}} = \Phi \left( -\frac{\nu_i^{\text{agg}}}{\sqrt{\Sigma[\text{agg}]_{ii}}} \right) > 0
\]

\[
\frac{\partial \text{agg}_\text{jump}^{[i]}}{\partial \xi^{[i]}} = -p^{\text{agg}} \phi \left( -\frac{\nu_i^{\text{agg}}}{\sqrt{\Sigma[\text{agg}]_{ii}}} \right) \frac{1}{\sqrt{\Sigma[\text{agg}]_{ii}}} < 0
\]

\[
\frac{\partial \text{Brownian}^{[i]}}{\partial p^{[i]}} = \frac{\partial \text{Brownian}^{[i]}}{\partial p^{\text{agg}}} = -\Phi \left( -\frac{\nu_i^{\text{none}}}{\sqrt{\Sigma[\text{none}]_{ii}}} \right) < 0
\]

\[
\frac{\partial \text{Brownian}^{[i]}}{\partial \xi^{[i]}} = -p^{[-i]} \phi \left( -\frac{\nu_i^{[-i]}}{\sqrt{\Sigma[i^{[-i]}]_{ii}}} \right) \frac{1}{\sqrt{\Sigma[i^{[-i]}]_{ii}}} - (1 - p^{[1]} - p^{[2]} - p^{\text{agg}}) \phi \left( -\frac{\nu_i^{\text{none}}}{\sqrt{\Sigma[\text{none}]_{ii}}} \right) \frac{1}{\sqrt{\Sigma[\text{none}]_{ii}}} < 0
\]

where here I have used that \( p^{\text{none}} = 1 - p^{[1]} - p^{[2]} - p^{\text{agg}} \).

As for the harder spillover risk in equation \((C.3)\), let \( F^x(x_1, x_2) \equiv F(x_1, x_2; \nu^x, \Sigma^x) \). Subscripts denote partial derivatives in the first two arguments. Then

\[
\frac{\partial \text{spillover}^{[1]}}{\partial p^{[1]}} = \left( F^{[1]}(\log(1 + \theta_{1,2}), 0) - F^{[1]}(0, 0) \right) - (F^{\text{none}}(\log(1 + \theta_{1,2}), 0) - F^{\text{none}}(0, 0)) \gtrless 0
\]

\[
\frac{\partial \text{spillover}^{[1]}}{\partial p^{[2]}} = \left( F^{[2]}(\log(1 + \theta_{1,2}), 0) - F^{[2]}(0, 0) \right) - (F^{\text{none}}(\log(1 + \theta_{1,2}), 0) - F^{\text{none}}(0, 0)) \gtrless 0
\]

\[
\frac{\partial \text{spillover}^{[1]}}{\partial p^{\text{agg}}} = \left( F^{\text{agg}}(\log(1 + \theta_{1,2}), 0) - F^{\text{agg}}(0, 0) \right) - (F^{\text{none}}(\log(1 + \theta_{1,2}), 0) - F^{\text{none}}(0, 0)) \gtrless 0
\]

\[
\frac{\partial \text{spillover}^{[1]}}{\partial \xi^{[1]}} = \sum_{x \in \{[1],[2],\text{agg, none}\}} p^x (-F^x_1(\log(1 + \theta_{1,2}), 0) + F^x_1(0, 0)) \gtrless 0
\]

\[
\frac{\partial \text{spillover}^{[1]}}{\partial \xi^{[2]}} = \sum_{x \in \{[1],[2],\text{agg, none}\}} p^x (-F^x_2(\log(1 + \theta_{1,2}), 0) + F^x_2(0, 0)) < 0
\]

The sign of \( \frac{\partial \text{spillover}^{[1]}}{\partial \xi^{[2]}} \) follows from each of the terms in the summation being positive.

Next, I match partial derivatives to obtain a linearization of spillover\(^{[1]}\) in the following five variables: idio\(_\text{jump}^{[1]}\), idio\(_\text{jump}^{[2]}\), p\(_\text{agg}\), Brownian\(^{[1]}\), and Brownian\(^{[2]}\). The sparsity the partials
gives us that we can solve for two coefficients at a time. This yields

\[ \text{spillover}^{[1]} \approx \left( \frac{\partial \text{idio\_jump}^{[1]} \partial \text{spillover}^{[1]}}{\partial p^{[1]} \partial \xi^{[1]}} - \frac{\partial \text{idio\_jump}^{[1]} \partial \text{spillover}^{[1]}}{\partial \xi^{[1]} \partial p^{[1]}} \right) \text{Brownian}^{[1]} + \left( \frac{\partial \text{idio\_jump}^{[2]} \partial \text{spillover}^{[1]}}{\partial p^{[2]} \partial \xi^{[2]}} - \frac{\partial \text{idio\_jump}^{[2]} \partial \text{spillover}^{[1]}}{\partial \xi^{[2]} \partial p^{[2]}} \right) \text{Brownian}^{[2]} + \left( \frac{\partial \text{idio\_jump}^{[2]} \partial \text{spillover}^{[2]}}{\partial p^{[2]} \partial \xi^{[2]}} - \frac{\partial \text{idio\_jump}^{[2]} \partial \text{spillover}^{[2]}}{\partial \xi^{[2]} \partial p^{[2]}} \right) \text{idio\_jump}^{[1]} + \left( \frac{\partial \text{idio\_jump}^{[2]} \partial \text{spillover}^{[2]}}{\partial p^{[2]} \partial \xi^{[2]}} - \frac{\partial \text{idio\_jump}^{[2]} \partial \text{spillover}^{[2]}}{\partial \xi^{[2]} \partial p^{[2]}} \right) \text{idio\_jump}^{[2]} + (\cdots) p^{agg} \]

where here I have suppressed the constant term. I have also suppressed the coefficient on \( p^{agg} \) as this coefficient is particularly messy and is not informative.\(^{22}\)

The denominators for the first four coefficients are all positive, as all the partial derivatives involved are negative except \( \frac{\partial \text{spillover}^{[1]}}{\partial p^{[1]}} \). Therefore, the signs of the coefficients depend on the signs of the partial derivatives of the spillover.

### C.5 Signs of Coefficients in Linearization

The coefficients in equation (C.4) are fairly opaque. I show that, under reasonable conditions, the coefficients on the cross-terms are positive, i.e., an increase in a default probability at bank 2 leads to a higher probability of a spillover at bank 1. The coefficients the bank’s own default probabilities are ambiguous.

#### C.5.1 Signs on Cross-Coefficients

Of the two relevant spillover-related partial derivatives, the only sign we know for sure is that \( \frac{\partial \text{spillover}^{[1]}}{\partial \xi^{[2]}} < 0 \). This one is also intuitive; whenever bank 2 has a larger asset buffer, it is less

\(^{22}\)The coefficient is simply \( \frac{\partial \text{spillover}^{[1]}}{\partial p^{agg}} \) minus each of the other four coefficients times their respective partial derivatives.
likely to default, and therefore also less likely to generate a spillover to bank 1. To following
proposition gives us sufficient conditions for the sign of \( \frac{\partial \text{spillover}^1}{\partial p^2} \) to be positive as well.

**Proposition C.1.** The following two conditions together are sufficient for \( \frac{\partial \text{spillover}^1}{\partial p^2} \geq 0 \):

(i) \( \Sigma_{12} \geq 0 \).

(ii) The mean of the jump is sufficiently negative:

\[
\mu_j^2 \leq \left( \sqrt{\frac{\sum_{22}^{\text{none}} - \frac{(\Sigma_{12}^{\text{none}})^2}{\Sigma_{11}^{\text{none}}}}{\Sigma_{22}^{\text{none}} - \frac{(\Sigma_{12}^{\text{none}})^2}{\Sigma_{11}^{\text{none}}}} - 1} \right) \left( \nu_2^{\text{none}} - \frac{\Sigma_{12}^{\text{none}}}{\Sigma_{11}^{\text{none}}} \nu_1^{\text{none}} \right)
\]

*Proof.* First, note that the marginal distribution of \( \log A_{\tau D}^1 \), prior to any spillovers, is the same conditional on no jump and conditional on an idiosyncratic jump at bank 2. Therefore, a sufficient condition for

\[
\Pr^{\text{none}}[1 < A_{\tau D}^1 \leq 1 + \theta_{1,2}, A_{\tau D}^2 \leq 1] \leq \Pr^{[2]}[1 < A_{\tau D}^1 \leq 1 + \theta_{1,2}, A_{\tau D}^2 \leq 1]
\]

is that the conditional probabilities satisfy

\[
\Pr^{\text{none}}[A_{\tau D}^2 \leq 1 \mid A_{\tau D}^1] \leq \Pr^{[2]}[A_{\tau D}^2 \leq 1 \mid A_{\tau D}^1]
\]

for all \( A_{\tau D}^1 \in (0, 1 + \theta_{1,2}) \).

These two conditional probabilities are given by

\[
\Pr^{\text{none}}[A_{\tau D}^2 \leq 1 \mid A_{\tau D}^1] = \Phi \left( -\frac{\nu_2^{\text{none}} + \frac{\Sigma_{12}^{\text{none}}}{\Sigma_{11}^{\text{none}}}(\log A_{\tau D}^1 - \nu_1^{\text{none}})}{\sqrt{\sum_{22}^{\text{none}} - \frac{(\Sigma_{12}^{\text{none}})^2}{\Sigma_{11}^{\text{none}}}}} \right)
\]

\[
\Pr^{[2]}[A_{\tau D}^2 \leq 1 \mid A_{\tau D}^1] = \Phi \left( -\frac{\nu_2^2 + \frac{\Sigma_{12}^2}{\Sigma_{11}^2}(\log A_{\tau D}^1 - \nu_1^2)}{\sqrt{\sum_{22}^{\text{none}} - \frac{(\Sigma_{12}^{\text{none}})^2}{\Sigma_{11}^{\text{none}}}} - \frac{(\Sigma_{12}^{\text{none}})^2}{\Sigma_{11}^{\text{none}}}} \right)
\]

\[
= \Phi \left( -\frac{\nu_2^2 + \frac{\Sigma_{12}^2}{\Sigma_{11}^2}(\log A_{\tau D}^1 - \nu_1^2)}{\sqrt{\sum_{22}^{\text{none}} + (\sigma_j^2)^2 - \frac{(\Sigma_{12}^{\text{none}})^2}{\Sigma_{11}^{\text{none}}}}} \right)
\]

Condition (ii) gives us that the desired inequality holds at \( A_{\tau D}^1 = 1 \). Condition (i) then gives us that it holds for all \( A_{\tau D}^1 \geq 1 \).

There are several comments about Proposition C.1 worth noting. First, the inequality condition in (ii) is effectively a limit on the probability of a positive jump. When \( (\sigma_j^2)^2 = 0 \), the condition just says that \( \mu_j \leq 0 \). As \( (\sigma_j^2)^2 \) increases, the bound on \( \mu_j^2 \) becomes larger in absolute value.
Second, there is an analogous lemma when $\Sigma_{12} < 0$. The derivation is the same, and simply requires ensuring that the inequality in the proof holds at $A^{[1]}_{\tau_D} = 1 + \theta_{1,2}$ instead of 1.

Third, the conditions are sufficient, but not necessary. $\mu_{\mathcal{J}}^{[2]}$ is allowed to be larger, especially if the majority of the mass of the marginal distribution of bank 1 prior to any spillovers lies well above 1.

Fourth, the expression

$$\nu_2^{\text{none}} - \frac{\Sigma_{12}^{\text{none}}}{\Sigma_{11}^{\text{none}}} \nu_1^{\text{none}} = \mathbb{E}_0^{\text{none}} [\log A^{[2]}_{\tau_D} | A^{[1]}_{\tau_D} = 1]$$

Therefore, given that $\mu_{\mathcal{J}}^{[2]} \leq 0$ was already assumed, condition (ii) in Proposition C.1 only becomes restrictive when bank 2 is expected to be in default when bank 1 hits its default boundary.

To understand the proposition, consider the following example of when it fails. Suppose $\Sigma_{12}^{\text{none}}$ is close to $\sqrt{\Sigma_{11}^{\text{none}} \Sigma_{22}^{\text{none}}}$, i.e., the two banks have little exposure to idiosyncratic diffusive risk, and therefore have a high positive correlation. Suppose furthermore that $A^{[1]}_0$ is close to 1 + $\theta_{1,2}$, and $A^{[2]}_0$ is close to 1, and that drifts are close to zero. Then, we have

$$\log A^{[1]}_{\tau_D} \approx \log (1 + \theta) + r_{\tau_D}$$
$$\log A^{[2]}_{\tau_D} \approx r_{\tau_D} + \text{jump}$$

where $r_{\tau_D}$ is the log-return on the aggregate asset they are both exposed to. Then, conditional on no jump, a spillover default occurs at bank 1 if $r_{\tau_D} < 0$. Conditional on a jump, a spillover default occurs if $r_{\tau_D} < 0$ and the jump is less than $-r_{\tau_D}$. This gives us that the probability of a spillover is lower conditional on a jump than conditional on no jump. Conditional on a jump, bank 2 is more likely to default, but these new defaults happen precisely when bank 1 is strong enough to survive the spillover. However, as long as $(\sigma_{\mathcal{J}}^{[2]})^2$ is positive, there is a nonzero chance of a positive jump. If $r_{\tau_D} < 0$, but the jump is bigger than $-r_{\tau_D}$, bank 2 survives and hence a spillover no longer happens.

With this sign restriction, we can make progress towards the signs of the coefficients in equation (C.4). We are then guaranteed that the coefficient on idio jmp$^{[2]}$ is positive.

**Corollary C.2.** The assumptions of Proposition C.1 are sufficient for obtaining a positive coefficient on idio jmp$^{[2]}$ in the linear expansion equation (C.4).

Even under these assumptions, the sign of the coefficient on Brownian$^{[2]}$ could be negative. If $\frac{\partial \text{idio jmp}}{\partial \xi^{[2]}} \ll 0$, then the coefficient becomes negative as well.

As before, these are sufficient, but not necessary, conditions. $\frac{\partial \text{spillover}^{[1]}}{\partial p^{[2]}}$ may be positive even if the conditions in Proposition C.1 fail, which would still give us unambiguously positive coefficients. In fact, even a mildly negative $\frac{\partial \text{spillover}^{[1]}}{\partial p^{[2]}}$ would still give us positive coefficients.
C.5.2 Signs on Own Coefficients

The coefficients in the linearization (C.4) on Brownian\[1\] and idio_jump[1] have ambiguous signs. The overall signs depend on where the majority of the mass lies for the distribution of $A_{\tau|D}^{1}\ | A_{\tau|D}^{2} \leq 1$ both conditional on a jump at bank 1 and conditional on no jump. There are

First, consider the sign of $\frac{\partial \text{spillover}^{[1]}}{\partial p^{[1]}}$. This derivative is positive if an idiosyncratic jump at 1 tends to push assets from above $1 + \theta_{1,2}$ down into the range $(1, 1 + \theta_{1,2}]$ conditional on a default at bank 2. In other words, it is positive if the jump makes bank 1 more sensitive to spillovers. The derivative is negative if an idiosyncratic jump at 1 tends to push assets from in the range $(1, 1 + \theta_{1,2}]$ down to below 1 conditional on a default at bank 2. Put differently, it is negative if the jump tends to cause bank 1 to default on its own before the spillovers are realized. The sign can be thought of as reflecting the balance between sensitivity and a race.

Next, consider the sign of $\frac{\partial \text{spillover}^{[1]}}{\partial \xi^{[1]}}$. This derivative is positive if, averaging over the four jump-cases, the increased buffer tends to push assets at bank 1 up into the range $(1, 1 + \theta_{2,1}]$ conditional on a default at bank 2, i.e., if bank 1 is no longer likely to default on its own before a spillover. The derivative is negative if, averaging over the four jump-cases, the increased buffer tends to push assets at bank 1 up above $1 + \theta_{2,1}$ conditional on a default at bank 2, i.e., if bank 1 is no longer sensitive to spillovers. Again, the sign can be thought of as reflecting the balance between sensitivity and a race.

There are effectively six cases to consider, each corresponding to a different level of assets at bank 1 and a different expected jump size at bank 1.

1. Bank 1 has a high level of assets, and idiosyncratic jumps are small enough that it tends to stay above the susceptible region. In this case, the
2. Bank 1 has a high level of assets, but idiosyncratic jumps put it in a region where it is susceptible to spillovers.
3. Bank 1 has a high level of assets, but idiosyncratic jumps are large enough to bankrupt bank 1 before any spillovers.
4. Bank 1 has a medium level of assets,
5. Bank 1 has a small level of assets.

D Correction to $(\mathbb{I} - D)^{-1}(C + f)$

As discussed in Section 2.3, the term $(\mathbb{I} - D)^{-1}C$ in equation (6) can contain elements that are a “mistake” from linearization. For example if $D_{i,j} > 0$ and $D_{j,i} > 0$, the multiplier $(\mathbb{I} - D)^{-1}$ would
imply that a shock to to bank \( i \) gets multiplied by passing through \( j \) and then coming back to \( i \). Obviously, the correct model would not have this feature: a default that originates at bank \( i \) in one state of the world cannot cause bank \( i \) to default in another state of the world.

Similarly, \((I - D)^{-1}f\) in equation (6) contains elements that appear as self-amplification. If \( \text{diag}((I - D)^{-1}) \) has \( i \)th entry not equal to 1, this is again saying that a correction to the default probability at \( i \) is amplified because any additional default at bank \( i \) causes bank \( i \) to default in more states of the world.

To be more precise, consider the following modification to the linear setup implied by equation (4). We explicitly consider all paths through which defaults propagate, and linearize these paths directly. As before, let \( c_{ij}^{[v]} \) denote the relative increase in the spillover default probability at bank \( i \) directly caused by the increase in the non-spillover default at bank \( j \), and let \( d_{ij}^{[v]} \) denote the relative increase in the spillover default probability at bank \( i \) caused by the increase in the spillover default at bank \( j \). Unlike before, let \( f_{ij}^{[v]} \) denote the relative adjustment to the spillover default probability at bank \( i \) to the non-spillover default probability at bank \( i \), but only representing the correction for the spillovers arising from bank \( j \). Roughly speaking, \( f^{[v]} \) in the setup of Section 2.3 can be thought of as \( \sum_j f_{ij}^{[v]} \). As before, this coefficient may be positive or negative depending on whether sensitivity or a race condition dominates, respectively.

The following diagram illustrates the setup.

\[
\begin{align*}
\Delta_{\text{non-spillover}}^{[v_1]}_{t+1} & \rightarrow c_{v_1}^{[v_2]} \rightarrow \Delta_{\text{spillover}}^{[v_2]}_{t+1} \rightarrow d_{v_2}^{[v_3]} \rightarrow \Delta_{\text{spillover}}^{[v_3]}_{t+1} \rightarrow \ldots \rightarrow d_{v_{m-1}}^{[v_m]} \rightarrow \Delta_{\text{spillover}}^{[v_m]}_{t+1} \\
\Delta_{\text{non-spillover}}^{[v_2]}_{t+1} & \rightarrow f_{v_2}^{[v_1]} \rightarrow \Delta_{\text{spillover}}^{[v_1]}_{t+1} \\
\end{align*}
\]

I start by considering a particular path \((v_1, \ldots, v_m)\) between banks that does not contain a cycle, i.e., \( 1 \leq v_1, \ldots, v_m \leq N \) for some \( m \geq 2 \) with \( v_1, \ldots, v_m \) all distinct. Bank \( v_1 \) affects the spillover default probability at bank \( v_2 \) in two ways. First, a change \( \Delta_{\text{non-spillover}}^{[v_1]}_{t+1} \) leads to a change \( c_{v_1}^{[v_2]} \Delta_{\text{non-spillover}}^{[v_1]}_{t+1} \) at bank \( v_2 \). Second, a change \( \Delta_{\text{non-spillover}}^{[v_2]}_{t+1} \) at bank \( v_2 \) affects the probability of a spillover default at bank \( v_2 \) arising from \( v_1 \) by \( f_{v_1}^{[v_2]} \Delta_{\text{non-spillover}}^{[v_2]}_{t+1} \). These two changes to the spillover then continue to propagate to bank \( v_3 \), with multiplicative coefficient \( d_{v_2}^{[v_3]} \), and so on.

Now, consider a particular bank \( i \). Given all the shocks \( \Delta_{\text{non-spillover}}^{[j]}_{t+1} \) at all the other banks \( j \neq i \), the total effect on bank \( i \) is obtained by summing over all paths that have \( i \) as its terminal bank. This include all paths of length 2 (the direct effects) that end at bank \( i \), as well as all paths of length 3 and longer (the indirect effects) that end at bank \( i \).

Mathematically, let \( P \) denote the set of all paths between banks that do not contain a cycle,
including the trivial path

\[ P = \{(v_1, \ldots, v_m) \mid m \geq 1 \land 1 \leq v_1, \ldots, v_m \leq N \land v_1, \ldots, v_m \text{ all distinct}\} \]

For any \(1 \leq i, j \leq N\), let \(P_{j \rightarrow i}\) denote the set of all paths from \(j\) to \(i\) that do not contain any cycles, i.e.,

\[ P_{j \rightarrow i} = \{(v_1, \ldots, v_m) \in P \mid v_1 = j \land v_m = i\} \]

Then

\[
\Delta_{\text{spillover}}^{[i]}_{t+1} = \sum_{j \neq i} \sum_{(v_1, \ldots, v_m) \in P_{j \rightarrow i}} \left( \prod_{k=3}^{m} d_{v_k}^{[v_k]} \right) (c_j^{[v_2]} \Delta_{\text{non-spillover}}^{[j]}_{t+1} + f_j^{[v_2]} \Delta_{\text{non-spillover}}^{[v_2]}_{t+1})
\]

To get this into a more interpretable form, for all \(i, j, k\) with \(i \neq k\) and \(j \neq k\), let \(P_{j \rightarrow i}^{[-k]}\) denote all paths from \(j\) to \(i\) that do not contain any cycles and that do not pass through \(k\), this time including the empty path, i.e.

\[ P_{j \rightarrow i}^{[-k]} = \{(v_1, \ldots, v_m) \in P \mid v_1 = j \land v_m = i \land k \notin \{v_1, \ldots, v_m\}\} \]

and let

\[
\tilde{C} = \left( \sum_{(v_1, \ldots, v_m) \in P_{j \rightarrow i}} c_j^{[v_2]} \prod_{k=3}^{m} d_{v_k}^{[v_k]} \right)_{i,j}
\]

\[
\tilde{f} = \left( \sum_{k \neq j} f_k^{[j]} \sum_{(v_1, \ldots, v_m) \in P_{j \rightarrow i}^{[-k]}} \prod_{\ell=2}^{m} d_{v_{\ell-1}}^{[v_\ell]} \right)_{i,j}
\]

Then

\[ \text{spillover}_t = (\tilde{C} + \tilde{f}) \text{ non-spillover}_t \]

Here, \(\tilde{C}\) is the corrected version of \((\mathbb{I} - D)^{-1}C\), and \(\tilde{f}\) is the corrected version of \((\mathbb{I} - D)^{-1}f\). Note that \(\text{diag}(\tilde{C}) = 0\) (there is no feedback to the originating bank), and \(\tilde{f}_{i,i} = \sum_{j \neq i} f_j^{[i]}\).

To make a comparison to the “uncorrected” coefficients, let \(\hat{P}_{j \rightarrow i}\) be all paths from \(j\) to \(i\), including the empty one and ones with cycles, and set \(f[i] = \sum_{j \neq i} f_j^{[i]}\). Then

\[
(\mathbb{I} - D)^{-1}C = \left( \sum_{(v_1, \ldots, v_m) \in \hat{P}_{j \rightarrow i}} c_j^{[v_2]} \prod_{k=3}^{m} d_{v_k}^{[v_k]} \right)_{i,j}
\]

\[
(\mathbb{I} - D)^{-1}f = \left( \sum_{k \neq j} f_k^{[j]} \sum_{(v_1, \ldots, v_m) \in \hat{P}_{j \rightarrow i}} \prod_{k=3}^{m} d_{v_{k-1}}^{[v_k]} \right)_{i,j}
\]

Comparing \(\tilde{C}\) to \((\mathbb{I} - D)^{-1}C\), we see the same structure except that we have reduced the number
of paths in the other sum to just those without cycles. Similarly, comparing \( \tilde{\mathbf{f}} \) to \((\mathbb{I} - \mathbf{D})^{-1}\mathbf{f}\), we see the same structure except that we have reduced the number of paths to eliminate any cycles.\(^{23}\) The corrected coefficients therefore still have the same interpretation. \( \tilde{\mathbf{C}}_{i,j} \) captures the change in bank \( i \)'s default probability in response to a change in the non-spillover default probability at bank \( j \) due to both the direct and indirect exposures it has to \( j \). \( \tilde{\mathbf{f}}_{i,j} \) captures the change in bank \( j \)'s role in transmitting shocks to bank \( i \) that originate elsewhere. The sum \((\tilde{\mathbf{C}} + \tilde{\mathbf{f}}_{i,j}) \) captures a mixture of 1) the direct transmission of defaults originating at bank \( j \) to bank \( i \), 2) the indirect transmission of defaults originating at bank \( j \), passing through some other banks, and then affecting \( i \), and 3) the changing role that bank \( j \) has in transmitting shocks from other banks to bank \( i \).

Note that the magnitudes for these coefficients may be drastically different. If \( \mathbf{D} \) has eigenvalues close to 1, the effect of truncation in the sum can be large. We can write \((\tilde{\mathbf{C}} + \tilde{\mathbf{f}}_{i,j}) \) as \( \tilde{\mathbf{D}}_{i,j} \) where \( \tilde{\mathbf{D}}_{i,j} \) is the corrected version that removes the originating bank in further transmissions. These defaults have to originate somewhere, and the sum in the corrected version also removes the originating bank in further transmissions.

### E Derivations and Proofs

#### E.1 Asset Prices

**Options Prices** Asset prices are simple to calculate in this setup. Note that conditional on know the ex-post number of jumps, equation (9) implies that the distribution of \( A_{t+\tau_i}^{[i]} \) is log-normal

\[
A_{t+\tau_i}^{[i]} | A_t^{[i]}, N_t^{[i]}(t + \tau_i) - N_t^{[i]}(t) = n \\
\sim \log N \left( \log A_t^{[i]} + \left( r_t - \frac{1}{2}(\sigma_t^{[i]})^2 \right) \tau_i + n\mu_{[j]}^{[i]}, (\sigma_t^{[i]})^2 \tau_i + n(\sigma_{[j]}^{[i]})^2 \right)
\]

Analogous to the usual Black and Scholes (1973) formulas, define for any \( K > 0 \)

\[
d_1^{[i]}(K) = \log \left( \frac{A_t^{[i]}}{K} + n\mu_{[j]} + (r_t + \lambda_t^{[i]}\xi_t^{[i]}) \right) - \frac{1}{2}(\sigma_t^{[i]})^2 \tau_i + n(\sigma_{[j]}^{[i]})^2 \\
d_2^{[i]}(K) = \log \left( \frac{A_t^{[i]}}{K} + n\mu_{[j]} + (r_t + \lambda_t^{[i]}\xi_t^{[i]}) - \frac{1}{2}(\sigma_t^{[i]})^2 \tau_i \right) \\
\]

With these, the price of call option on bank \( i \) at time \( t \) with strike \( k \) and maturity \( \tau_i \) conditional on \( n \) jumps occurring is

\[
C_i^{[i]}(k, n, \tau_i) = A_t^{[i]} e^{\frac{r_t^{[i]}\tau_i}{2} + n\left( \mu_{[j]}^{[i]} + \frac{1}{2}(\sigma_{[j]}^{[i]})^2 \right)} \Phi(d_1^{[i]}) - e^{-r_t^{[i]}\tau_i} k \Phi(d_2^{[i]})
\]

\(^{23}\)The reduction in cycles here goes one step further. This part of the coefficient is about bank \( j \)'s role in transmitting defaults further down the line. These defaults have to originate somewhere, and the sum in the corrected version also removes the originating bank in further transmissions.
where \(\Phi(\cdot)\) is the normal CDF. This is just the usual Black-Scholes call price adjusted for \(n\) jumps occurring.

Using that the number of jumps is an independent Poisson distribution, we get that the price of a call option on bank \(i\) at time \(t\) with strike \(k\) and maturity \(\tau^D\) is

\[
C_t^{[i]}(k, \tau^D) = \sum_{n=0}^{\infty} \frac{(\lambda_t^{[i]} \tau^D)^n e^{-\lambda_t^{[i]} \tau^D}}{n!} C_t^{[i]}(k, n, \tau^D)
\]

As \(\lambda_t^{[i]} \tau^D\) is small, I compute this sum by truncating the sum, and putting all of the remaining weight in the last term.

A put option can be computed with the usual put-call parity formula

\[
P_t^{[i]}(k, \tau^D) = C_t^{[i]}(k, \tau^D) + e^{-r \tau^D} k - A_t^{[i]}
\]

Finally, the probability of remaining above a certain value \(k\) is simply

\[
\Pr_t^Q \left[ A_{t+\tau^D} > k \right] = \sum_{n=0}^{\infty} \frac{(\lambda_t^{[i]} \tau^D)^n e^{-\lambda_t^{[i]} \tau^D}}{n!} \Phi(-d_2^{[i]})
\]

which can again be approximated by truncating the sum.

**Black-Scholes Implied Volatilities** For both put and call options, I compute these using the traditional numerical methods. I use the QuantLib library through the RQuantLib package in R to perform these computations in the data. For the estimator, when computing the model-implied BSIVs, I use a standard Newton method approach.

Although the values have no closed form solution, the Implicit Function Theorem does yield analytical formulas for the derivatives in terms of the BSIVs. Using that the model-implied call and put values have closed form derivatives, I use this to compute the derivative and Hessian of the objective (10). This greatly improves performance, especially since the minimization problem is quite sparse. State variables on different days do not interact with each other, and only interact indirectly through the parameters.

**E.2 Proof of Proposition 1**

**Proof.** For notational convenience, let \(\Gamma = (I - D)^{-1} (C + f)\). Row \(i\) of equation 6 implies that

\[
\Delta \text{jump,like}_{t+1}^{[i]} = \Gamma_{i,:} \Delta \text{Brownian}_{t+1} + (e_i + \Gamma) \text{agg \ jump}_{t+1} + (e_i + \Gamma_i,:) \text{idio \ jump}_{t+1}
\]

where here \(e_i\) is the \(i\)th coordinate vector. From the Frisch-Waugh-Lovell theorem, we know that estimating \(\gamma_{j}^{[i]}\) in regression equation (7) is equivalent to the univariate regression after first
conditioning linearly on the remaining variables. Hence
\[
\gamma[i]_j = \frac{\text{cov}(\Delta\text{jump}_{\text{like}}[i]_{t+1}, \Delta\text{Brownian}[j]_{t+1} | \Delta\text{Brownian}[-j]_{t+1}, \Delta\text{agg \_jump}_{t+1})}{\text{var}(\Delta\text{Brownian}[j]_{t+1} | \Delta\text{Brownian}[i]_{t+1}, \Delta\text{agg \_jump}_{t+1})}
\]
where as in Assumption 2, the conditional covariance and conditional variance are interpreted as the covariance and variance after conditioning on the conditional variables linearly.

Note that for \(k \neq j\) we have
\[
\text{cov}(\Delta\text{Brownian}[k]_{t+1}, \Delta\text{Brownian}[j]_{t+1} | \Delta\text{Brownian}[-j]_{t+1}, \Delta\text{agg \_jump}_{t+1}) = 0
\]
as \(\Delta\text{Brownian}[k]_{t+1}\) is one of the conditioning variables. Similarly, for all \(k\) (including \(j\))
\[
\text{cov}(\Delta\text{agg \_jump}_{t+1}, \Delta\text{Brownian}[j]_{t+1} | \Delta\text{Brownian}[-j]_{t+1}, \Delta\text{agg \_jump}_{t+1}) = 0
\]
due to the presence of \(\Delta\text{agg \_jump}_{t+1}\) in the conditioning variables. Hence, for \(j \neq i\),
\[
\gamma[i]_j = \Gamma_{i,j} + \sum_k (e_i + \Gamma_{i,:})_k \frac{\text{cov}(\Delta\text{idio \_jump}[k]_{t+1}, \Delta\text{Brownian}[j]_{t+1} | \Delta\text{Brownian}[-j]_{t+1}, \Delta\text{agg \_jump}_{t+1})}{\text{var}(\Delta\text{Brownian}[j]_{t+1} | \Delta\text{Brownian}[-j]_{t+1}, \Delta\text{agg \_jump}_{t+1})}
\]
Assumption 2 can be rewritten as a regression
\[
\Delta\text{idio \_jump}[k]_{t+1} = \rho[k] \Delta\text{Brownian}[k]_{t+1} + \omega[k] \Delta\text{agg \_jump}_{t+1} + v[k]_{t+1}
\]
where \(\text{cov}(v[k]_{t+1}, \Delta\text{Brownian}[\ell]_{t+1}) = 0\) for all \(\ell\), including \(k\), and \(\text{cov}(v[k]_{t+1}, \Delta\text{agg \_jump}_{t+1}) = 0\). Hence, for \(k \neq j\),
\[
\text{cov}(\Delta\text{idio \_jump}[k]_{t+1}, \Delta\text{Brownian}[j]_{t+1} | \Delta\text{Brownian}[-j]_{t+1}, \Delta\text{agg \_jump}_{t+1}) = 0
\]
For \(k = j\),
\[
\text{cov}(\Delta\text{idio \_jump}[j]_{t+1}, \Delta\text{Brownian}[j]_{t+1} | \Delta\text{Brownian}[-j]_{t+1}, \Delta\text{agg \_jump}_{t+1}) = \rho[j] \text{var}(\Delta\text{Brownian}[j]_{t+1} | \Delta\text{Brownian}[-j]_{t+1}, \Delta\text{agg \_jump}_{t+1})
\]
Hence, since \(j \neq i\), we get that
\[
\gamma[i]_j = (1 - \rho[j])\Gamma_{i,j}
\]
The proposition is proven by noting that

\[ \rho^j = \frac{\text{cov} \left( \Delta \text{idio\_jump}^j_{t+1}, \Delta \text{Brownian}^j_{t+1} \mid \Delta \text{agg\_jump}_{t+1} \right)}{\text{var} \left( \Delta \text{Brownian}^j_{t+1} \mid \Delta \text{agg\_jump}_{t+1} \right)} \]

\[ \square \]

E.3 Reverse Causality Coefficient in Equation (12)

Note that

\[ \tilde{\Delta \text{Brownian}}^2_{t+1} = \tilde{\Delta \text{Brownian\_own}}^2_{t+1} + \alpha(\tilde{\Delta \text{agg\_jump}}^1_{t+1} + \tilde{\Delta \text{idio\_jump}}^1_{t+1}) \]

The tilde indicates orthogonalizing out \( \Delta \text{agg\_jump}_{t+1} \) and \( \Delta \text{Brownian}^1_{t+1} \), so this becomes

\[ \Delta \text{Brownian}^2_{t+1} = \Delta \text{Brownian\_own}^2_{t+1} + \alpha \tilde{\Delta \text{idio\_jump}}^1_{t+1} \]

Hence

\[ \text{cov}(\tilde{\Delta \text{jump\_like}}^1_{t+1}, \Delta \text{Brownian}^2_{t+1}) = \alpha \text{var}(\tilde{\Delta \text{idio\_jump}}^1_{t+1}) \]

\[ \text{var}(\Delta \text{Brownian}^2_{t+1}) = \text{var}(\Delta \text{Brownian\_own}^2_{t+1}) + \alpha^2 \text{var}(\tilde{\Delta \text{idio\_jump}}^1_{t+1}) \]

Taking the ratio and rearranging yields \( \gamma_2^1 \) in equation (12).

F More Simulations

Describe the simulated tables.
Table F.1: Regression results in simulated data.

Panel (a) shows the parameter values and initial conditions used. 1,000 simulations were drawn, each with two years (504 trading days) worth of data. Draws resulting in a realized default are discarded. On each day, default probabilities are computed at the six month (126 trading day) horizon. Each default probability is computed as a Monte-Carlo average of 1,000 draws.

Panels (b) show the results of running regression equation (7) on of the generated data. Columns correspond to different left-hand-side variables, while rows correspond to right-hand-side variables. Non-zero values are evidence of a direct or indirect linkage. Recall:

\[
\Delta \text{jump}_i[t+1] = \alpha[i] + \beta[i] \Delta \text{aggjump}_i[t+1] + \chi[i] \Delta \text{Brownian}_i[t+1] + \sum_{j \neq i} \gamma[i][j] \Delta \text{Brownian}_j[t+1] + u_i[t+1]
\]

(a) Parameter values and initial conditions.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Volatility</td>
<td>(\sigma)</td>
<td>Bank Volatility</td>
<td>(\sigma)</td>
</tr>
<tr>
<td>Aggregate Jump</td>
<td>(\kappa, \lambda, \sigma_{\lambda})</td>
<td>Bank Jump</td>
<td>(\kappa)</td>
</tr>
<tr>
<td></td>
<td>(-0.20, 0.10)</td>
<td>(\lambda)</td>
<td>((0.02, 0.02, 0.02))</td>
</tr>
<tr>
<td>Bank Loadings</td>
<td>(\beta_B)</td>
<td>Initial Conditions</td>
<td>(\lambda_0)</td>
</tr>
<tr>
<td></td>
<td>((1.00, 1.00, 1.00))</td>
<td></td>
<td>(\lambda_0)</td>
</tr>
<tr>
<td></td>
<td>((1.00, 1.00, 1.00))</td>
<td></td>
<td>(\Theta)</td>
</tr>
<tr>
<td>Spillovers</td>
<td>(\Theta)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>((0.00, 0.20, 0.00))</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>((0.00, 0.00, 0.00))</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) Regression (7).

<table>
<thead>
<tr>
<th>Jump</th>
<th>bank1</th>
<th>bank2</th>
<th>bank3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brownian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aggregate jump</td>
<td>0.13</td>
<td>0.18</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(1.31)</td>
<td>(1.04)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>bank1</td>
<td>-0.93</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.04)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>bank2</td>
<td>0.15</td>
<td>-0.82</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.22)</td>
<td>(0.27)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>bank3</td>
<td>0.63</td>
<td>0.72</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.28)</td>
<td>(0.29)</td>
<td>(0.02)</td>
</tr>
</tbody>
</table>
Table F.2: Regression results in simulated data.

Panel (a) shows the parameter values and initial conditions used. 1,000 simulations were drawn, each with two years (504 trading days) worth of data. Draws resulting in a realized default are discarded. On each day, default probabilities are computed at the six month (126 trading day) horizon. Each default probability is computed as a Monte-Carlo average of 1,000 draws.

Panels (b) show the results of running regression equation (7) on the generated data. Columns correspond to different left-hand-side variables, while rows correspond to right-hand-side variables. Non-zero values are evidence of a direct or indirect linkage. Recall:

\[
\Delta \text{jump}_{i, t+1} = \alpha_i + \beta_i \Delta \text{agg jump}_{t+1} + \chi_i \Delta \text{Brownian}_{i, t+1} + \sum_{j \neq i} \gamma_j \Delta \text{Brownian}_{j, t+1} + u_{i, t+1}
\]

(a) Parameter values and initial conditions.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Volatility</td>
<td>( \sigma )</td>
<td>Bank Volatility</td>
<td>( \sigma )</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>0.10</td>
<td>(0.05, 0.05, 0.05)</td>
<td></td>
</tr>
<tr>
<td>Aggregate Jump</td>
<td>( \kappa, \lambda, \sigma_\lambda )</td>
<td>Bank Jump</td>
<td>( \kappa )</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>1.00, 0.05, 0.10</td>
<td>(1.00, 1.00, 1.00)</td>
<td></td>
</tr>
<tr>
<td>( \lambda )</td>
<td>-0.20, 0.10</td>
<td>(0.02, 0.02, 0.02)</td>
<td></td>
</tr>
<tr>
<td>( \sigma_\lambda )</td>
<td>0.10</td>
<td>(0.10, 0.10, 0.10)</td>
<td></td>
</tr>
<tr>
<td>Bank Loadings</td>
<td>( \beta_{bi} )</td>
<td>Initial Conditions</td>
<td>( \lambda_0 )</td>
</tr>
<tr>
<td>( \beta_{bi} )</td>
<td>(1.00, 1.00, 1.00)</td>
<td>( \lambda_0 )</td>
<td>0.05</td>
</tr>
<tr>
<td>( \beta_{bj} )</td>
<td>(1.00, 1.00, 1.00)</td>
<td>( \lambda_0 )</td>
<td>(1.15, 1.15, 1.15)</td>
</tr>
<tr>
<td>Spillovers</td>
<td>( \Theta )</td>
<td>( \lambda_0 )</td>
<td>(0.02, 0.02, 0.02)</td>
</tr>
</tbody>
</table>
| \( \Theta \)       | \[
\begin{pmatrix}
0.05 & 0.00 \\
0.00 & 0.05
\end{pmatrix}
\] |

(b) Regression (7).

<table>
<thead>
<tr>
<th>Jump</th>
<th>bank1</th>
<th>bank2</th>
<th>bank3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brownian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aggregate jump</td>
<td>0.17</td>
<td>0.17</td>
<td>0.19</td>
</tr>
<tr>
<td>bank1</td>
<td>-0.86</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>bank2</td>
<td>0.16</td>
<td>-0.76</td>
<td>0.00</td>
</tr>
<tr>
<td>bank3</td>
<td>0.61</td>
<td>0.67</td>
<td>-0.01</td>
</tr>
</tbody>
</table>
Table F.3: Regression results in simulated data.

Panel (a) shows the parameter values and initial conditions used. 1,000 simulations were drawn, each with two years (504 trading days) worth of data. Draws resulting in a realized default are discarded. On each day, default probabilities are computed at the six month (126 trading day) horizon. Each default probability is computed as a Monte-Carlo average of 1,000 draws.

Panels (b) show the results of running regression equation (7) on of the generated data. Columns correspond to different left-hand-side variables, while rows correspond to right-hand-side variables. Non-zero values are evidence of a direct or indirect linkage. Recall:

\[
\Delta \text{jump}_t = \alpha_i + \beta_i \Delta \text{agg\_jump}_t + \chi_i \Delta \text{Brownian}_t + \sum_{j \neq i} \gamma_{ij} \Delta \text{Brownian}_t + u_t
\]

(a) Parameter values and initial conditions.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Volatility σ</td>
<td>0.10</td>
<td>Bank Volatility σ</td>
<td>(0.15, 0.15, 0.15)</td>
</tr>
<tr>
<td>Aggregate Jump κ, λ, σλ</td>
<td>1.00, 0.05, 0.05</td>
<td>Bank Jump μ</td>
<td>(1.00, 1.00, 1.00)</td>
</tr>
<tr>
<td>µj, σj</td>
<td>-0.20, 0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bank Loadings αB, αJ</td>
<td>(0.00, 0.00, 0.00)</td>
<td>Initial Conditions λo</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>(1.00, 1.00, 1.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spillovers</td>
<td>Θ</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00, 0.05, 0.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00, 0.00)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) Regression (7).

<table>
<thead>
<tr>
<th>Jump</th>
<th>bank1</th>
<th>bank2</th>
<th>bank3</th>
</tr>
</thead>
<tbody>
<tr>
<td>aggregate jump</td>
<td>-0.17</td>
<td>0.09</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(1.97)</td>
<td>(1.72)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>bank1</td>
<td>-0.75</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.28)</td>
<td>(0.06)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>bank2</td>
<td>0.05</td>
<td>-0.68</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.33)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>bank3</td>
<td>0.04</td>
<td>0.19</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.23)</td>
<td>(0.01)</td>
</tr>
</tbody>
</table>
Table F.4: Regression results in simulated data.

Panel (a) shows the parameter values and initial conditions used. 1,000 simulations were drawn, each with two years (504 trading days) worth of data. Draws resulting in a realized default are discarded. On each day, default probabilities are computed at the six month (126 trading day) horizon. Each default probability is computed as a Monte-Carlo average of 1,000 draws.

Panels (b) show the results of running regression equation (7) on the generated data. Columns correspond to different left-hand-side variables, while rows correspond to right-hand-side variables. Non-zero values are evidence of a direct or indirect linkage. Recall:

\[
\Delta \text{jump}_{i,t+1} = \alpha_i + \beta_i \Delta \text{agg jump}_{t+1} + \chi_i \Delta \text{Brownian}_i + \sum_{j \neq i} \gamma_{ij} \Delta \text{Brownian}_j + \varepsilon_{i,t+1}
\]

(a) Parameter values and initial conditions.

(b) Regression (7).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Volatility</td>
<td>(\sigma = 0.10)</td>
<td>Bank Volatility</td>
<td>(\sigma = (0.05, 0.05, 0.05))</td>
</tr>
<tr>
<td>Aggregate Jump</td>
<td>(\kappa, \lambda, \sigma_\lambda = (1.00, 0.10, 0.10))</td>
<td>Bank Jump</td>
<td>(\chi = (1.00, 1.00, 1.00))</td>
</tr>
<tr>
<td>(\mu_j, \sigma_j)</td>
<td>(-0.20, 0.10)</td>
<td>(\sigma_\lambda)</td>
<td>(0.10, 0.10, 0.10)</td>
</tr>
<tr>
<td>Bank Loadings</td>
<td>(\beta_j = (1.00, 1.00, 1.00))</td>
<td>Initial Conditions</td>
<td>(\lambda_0 = 0.10)</td>
</tr>
</tbody>
</table>
| \(\Theta\)      | \[
\begin{bmatrix}
0.20 & 0.00 \\
0.00 & 0.20 \\
0.00 & 0.00
\end{bmatrix}
\] | \(\Lambda_0 = (1.15, 1.15, 1.15)\) |
|                 |           |                 | \(\varepsilon_0 = (0.02, 0.02, 0.02)\) |

<table>
<thead>
<tr>
<th>Jump</th>
<th>bank1</th>
<th>bank2</th>
<th>bank3</th>
</tr>
</thead>
<tbody>
<tr>
<td>aggregate jump</td>
<td>0.37</td>
<td>0.38</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>(0.55)</td>
<td>(0.44)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>bank1</td>
<td>-0.89</td>
<td>-0.02</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.07)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>bank2</td>
<td>0.24</td>
<td>-0.73</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.29)</td>
<td>(0.31)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>bank3</td>
<td>0.70</td>
<td>0.78</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.33)</td>
<td>(0.31)</td>
<td>(0.03)</td>
</tr>
</tbody>
</table>

Table F.5: Regression results in simulated data.

Panel (a) shows the parameter values and initial conditions used. 1,000 simulations were drawn, each with two years (504 trading days) worth of data. Draws resulting in a realized default are discarded. On each day, default probabilities are computed at the six month (126 trading day) horizon. Each default probability is computed as a Monte-Carlo average of 1,000 draws.

Panels (b) show the results of running regression equation (7) on of the generated data. Columns correspond to different left-hand-side variables, while rows correspond to right-hand-side variables. Non-zero values are evidence of a direct or indirect linkage. Recall:

\[
\Delta \text{jump}_i^{t+1} = \alpha_i^t + \beta_i^t \Delta \text{agg jump}_{t+1} + \chi_i^t \Delta \text{Brownian}_i^{t+1} + \sum_{j \neq i} \gamma_j^t \Delta \text{Brownian}_j^{t+1} + \epsilon_i^{t+1}
\]

(a) Parameter values and initial conditions.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Volatility</td>
<td>σ</td>
<td>Bank Volatility</td>
<td>σ</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>(0.05, 0.05, 0.05)</td>
<td></td>
</tr>
<tr>
<td>Aggregate Jump</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>κ, λ, σ_κ</td>
<td>1.00, 0.05, 0.10</td>
<td>κ</td>
</tr>
<tr>
<td></td>
<td>μ_J, σ_J</td>
<td>-0.20, 0.10</td>
<td>λ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>σ_λ</td>
</tr>
<tr>
<td>Bank Loadings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>β_B</td>
<td>(1.00, 1.00, 1.00)</td>
<td>λ_0</td>
</tr>
<tr>
<td></td>
<td>β_J</td>
<td>(1.00, 1.00, 1.00)</td>
<td>A_0</td>
</tr>
<tr>
<td>Spillovers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Θ</td>
<td>0.00 0.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.00 0.00</td>
<td></td>
</tr>
</tbody>
</table>

(b) Regression (7).

<table>
<thead>
<tr>
<th>Brownian</th>
<th>bank1</th>
<th>bank2</th>
<th>bank3</th>
</tr>
</thead>
<tbody>
<tr>
<td>aggregate jump</td>
<td>0.15</td>
<td>0.17</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>(0.74)</td>
<td>(0.59)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>bank1</td>
<td>-0.89</td>
<td>-0.02</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.07)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>bank2</td>
<td>0.24</td>
<td>-0.73</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.29)</td>
<td>(0.31)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>bank3</td>
<td>0.70</td>
<td>0.77</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.33)</td>
<td>(0.31)</td>
<td>(0.02)</td>
</tr>
</tbody>
</table>
GARCH Model of Holding Period Returns

The MA(125)-GJR-GARCH(1,1) model in Section 6.2 is as follows

\[ \log \text{eHPR}_t^{[i]} = \mu_t + \left( \sum_{i=1}^{125} z_{t-i} \right) + z_t \]  
\[ \mu_t = \mu + \chi \text{CDS}^{[i]}_{6m,t-6m} \]  
\[ z_t = \sigma_t \epsilon_t, \quad \epsilon_t \overset{iid}{\sim} N(0,1) \]  
\[ \sigma_{t+1}^2 = \omega + (\alpha + \gamma 1\{z_t \leq 0\}) z_t^2 + \beta \sigma_t^2 \]

As described in Section 6.2, the MA(125) mean model is designed to capture the substantial overlap in holding period returns between dates less than six months apart. For fixed volatility, the covariance between two dates falls linearly, until it reaches zero at observations at least six months apart.

The GJR-GARCH(1,1) model for the variance is based on Glosten et al. (1993). The parameters \( \omega, \alpha, \) and \( \beta \) are from the usual GARCH(1,1) specification. \( \alpha \) captures the sensitivity to realized shocks, while \( \beta \) allows for an autoregressive decay of the variances. The parameter \( \gamma \) allows for the response to realized shocks to be asymmetric. In particular, for \( \gamma > 0 \) the model allows for variances to increase more in response to negative shocks than positive shocks.

The mean model \( \mu_t \) in equation (G.1) loads linearly on the six month CDS from six months prior. This is a linear approximation to the true expectation of the excess holding period return. Unfortunately, the true expected return, conditional on no default, is not simply the short-term CDS spread. In most models, it is increasing in the CDS spread. The affine term in the mean model is meant to capture this dependence.

I fit the model in equation (G.1) using therugarch package in \( R \) by Ghalanos (2014). To compute the model-implied default probabilities, I use the fitted point estimates, and simulate out 10,000 paths of six months (126 days) of data. I then compute the probability of the event in (14) as a Monte-Carlo average over these draws of the indicator for

\[ \mu_{t+6m} + \left( \sum_{i=1}^{125} z_{t+6m-i} \right) + z_{t+6m} \leq -\log A \]

To compute first differences of the Brownian risk, I ensure that both probabilities of default are computed at the same horizon. For time \( t \), this simply means using the same value for \( \mu_t \), and the realized \( z_{t+1}, z_{t+2} \) and onward are simulated using the newly estimated \( \sigma_{t+2} \).