Global Corporate Default Clustering and Contagion*

Yanru Lee†

University of North Carolina at Chapel Hill

February 12, 2023

Abstract

This paper evaluates the severity of international corporate default clustering. By employing a dataset that contains a global coverage of corporate default events, I show that corporate default events display an excess degree of clustering in four key regions worldwide: U.S., Europe, Asia Pacific, and Emerging Markets. Specifically, I show that systemic risk drives the clustering of corporate default events, which cannot be explained by firm fundamentals variables and key systematic variables identified in corporate default risk literature. After controlling for the impact of domestic corporate default contagion within a region, I identify a novel source of default clustering: International corporate default contagion.

Keywords: Corporate Default Clustering, International Default Contagion, Systemic Risk
JEL Classifications: F3, G15, G33, C50

*I am highly indebted to Professor Anusha Chari for the excellent guidance and mentorship. I am also especially grateful for the helpful feedbacks from Professor Peter Hansen, Christian Lundblad, Elena Simintzi, Andrii Babii, Eric Ghysels, as well as seminar participants in the 17th Annual Conference on Asia-Pacific Financial Markets (CAFM), Inter-Finance PhD Seminar and UNC Chapel Hill. Preliminary work for this paper was written when Yanru Lee was a visiting scholar at Credit Research Initiative at National University of Singapore (NUS CRI). I am highly grateful to Professor Jin-Chuan Duan for facilitating my visiting appointment, and the team at NUS CRI for providing the requested data for my research. Without this dataset, this research would not have been possible. All errors are my own.

†Graduate Student, University of North Carolina at Chapel Hill. Email: leeyanru@live.unc.edu

1 Introduction

The global economy has undergone through several adverse episodes of corporate default clustering in the last few decades. Figure 1 depicts the monthly number of corporate default events pertaining to public firms from 1996 – 2020 in four key regions of the world: the U.S., Europe, Asia Pacific, and Emerging Markets. The Figure shows that two of the most severe global episodes of default clustering took place during the early 2000s and 2008 – 2009 period. These two periods coincide with the global macroeconomic recession in the early 2000s, and the global financial crisis period, respectively. Apart from those two episodes, corporate default data also exhibits several severe clusters that are confined at a regional scale. For instance, Europe suffered through clusters of default events during part of the Eurozone Sovereign Debt Crisis. Asia Pacific and Emerging Markets also went through several intermittent clustering of corporate default events during the late 1990s, which coincided with the Asian Financial Crisis and the default of Russian debt in 1998.

Corporate default clustering poses both risk and asset pricing implications. Several studies on default clustering in the U.S. economy point out that systemic risk plays a substantial role in corporate debt distress. Particularly, Das et al. (2007), Duffie et al. (2009), Giesecke and Kim (2011), and others show that firm-specific and systematic variables are inadequate in accounting for the degree of realized clustering of corporate default events, especially under crisis periods. These findings raise concerns over underestimating risks to financial stability. Specifically, an aggregate estimate of economy-level corporate default risk exposure, solely based on firm-specific and systematic variables, would have underestimated potential vulnerability in corporate debt due to the clustering of debt distress. From the asset pricing perspective, default clustering contributes to the distress risk premium puzzle. Chava and Purnanandam (2010) document empirical evidence that an unanticipated surge of U.S. corporate default events leads to an anomalous negative relation between equity returns and distress risk. Despite clear evidence on the severity of corporate default clustering worldwide, existing research in this field almost exclusively focuses on the U.S. economy, with no further analysis on the severity of default clustering in the rest of the world.

By employing a dataset that contains a global coverage of corporate default events, this paper evaluates the severity of corporate default clustering on an international scale. As corporate default events are rare, I focus on the severity of corporate default clustering at a regional scale in four separate regions: U.S. economy, Europe, Asia Pacific, and Emerging Markets. Based on this approach, this paper makes the following contributions:

To begin with, by applying the Time Rescaling Poisson Method, I show that corporate default events in the four economic regions display a severe degree of excess default clustering beyond the information contained in firm fundamentals and key systematic variables identified in corporate default risk literature.

Next, I show that systemic risk drives the excess clustering of default events, not due to severe random intermittent spikes of default events in a specific period. By measuring systemic risk as a serial dependence between extreme quantile waves of default events in a region, I show that a severe wave of default events in a period increases the risk of similar multiple adverse waves of default events for up to two years.

Finally, I identify significant evidence of international corporate default contagion as a novel source of clustering in corporate default events. Across the four regions, I separately estimate a Poisson Autoregression Model on corporate default events in time rescaled data. The results show significant evidence of persistent cross-border corporate default contagion at a regional scale. Impulse response analysis shows that an adverse shock of corporate default events in a region dynamically elevates the risk of additional waves of default events over multiple periods. The impact gradually decays over time. I subsequently extend the Poisson Autoregression Model to a multivariate setting and identify significant evidence of international cross-region corporate default contagion.

International trade, specifically literature on the global value chain, points out that multinational firms have incentives to develop extensive relationships with multiple firms on an international scale. The global network of firms' relationships is often complex, consisting of multiple layers of intricate networks among numerous firms in different continents. This relationship, however, also facilitates the failure or default by a critical firm (or firms) in fulfilling its financial obligations to potentially disrupt the global economy and spread financial distress risk to interconnected firms. In severe cases, the global financial system may suffer from a cascade of correlated default events beyond the information contained in firm-specific and systematic variables. This result is triggered, as explanatory variables may not adequately account for information on firms' cross-border business interlinkages and the corresponding impact of financial distress risk contagion.

International finance literature also suggests that corporate default events can be contagious across different economies and even globally. Financially troubled firms may face difficulties fulfilling their tax obligations, which increases the default risk of the respective country in which it domiciles. As different countries often borrow from common lenders, default by a single country may adversely impact other countries' propensity to obtain financial loans, potentially triggering default risk contagion across countries. In turn, countries that face elevated default risk may also signal a poor economic outlook. Firms operating within this economy face an elevated cost of raising funds, thus increasing their risk of financial distress.

This paper begins by using a proportional hazard model to measure individual firms' default risk exposure. Across the four key regions, I separately estimate the proportional hazard models. An important aspect of our paper lies in the comprehensive selection of the relevant explanatory variables based on existing corporate default risk literature.

In the process, I consider a wide array of firm-specific, domestic macroeconomics and global factors that are relevant in assessing a firm's default risk. This approach is crucial for evaluating whether corporate default events still display excess clustering, even after incorporating the information in a wide array of explanatory variables.

Another key distinction of this paper lies in including financial firms in the data sample. Existing corporate default clustering research based on firm-level data cannot justifiably include financial and non-financial firms in the same data sample. This is because financial and non-financial firms have distinct capital structure. Including both types of firms in the same data sample may distort overall corporate default risk assessment, forcing the exclusion of financial firms or study these two types of firms separately. However, this approach may inevitably lead to a misrepresented evaluation in the severity of default clustering due to the role and importance of financial firms in the overall economy.

I can justifiably include financial firms in the data sample without a noticeable decline in the overall corporate default risk assessment. I do so by incorporating a non-standard calculation of the distance to default (DtD) variable, following Duan et al. (2012), replacing firm leverage. DtD refers to the volatility-adjusted leverage and indicates a firm's default risk based on leverage and asset volatility. To justifiably include financial firms in the data sample, Duan et al. (2012) propose a non-standard calculation of DtD by introducing an additional parameter that adjusts for inherent differences in leverage ratio between financial firms and non-financial firms. I follow Duan et al. (2012) measurement of DtD to justifiably include financial firms in the data sample for assessing the severity of corporate default clustering.

Using my measure of corporate default risk, I apply the time rescaling method on the cumulative aggregate corporate default intensity at the regional level. In the context of corporate default risk assessment, Das et al. (2007) pioneer this method to evaluate if explanatory variables can sufficiently account for the clustering of default events in the U.S. economy. The time rescaling Poisson method rescales the time interval of accumulated aggregate corporate default intensity. The result is that corresponding actual corporate default events in rescaled time intervals should follow a Poisson distribution if the proportional hazard model is correctly specified. Plots that compares the corporate default data in time rescaled interval and its theoretical Poisson distribution counterpart shows that empirical default data exhibits heavy tail, and do not follow a Poisson distribution.

Following the empirical plots, I critically evaluate if corporate default events display an excess degree of clustering not explained by the information contained in explanatory variables. I apply a series of econometrics tests on the corporate default data in time rescaled intervals to do so. The econometrics tests jointly assess if default data follows a Poisson distribution. Specifically, the tests focus on the distributional, dispersion, tail properties, generating function, and serial dependence of the default data in the time rescaled interval. Across all four regions, I document that the default data in time rescaled interval largely fails all the econometrics tests, specifically showing strong evidence that default data in all four regions exhibit a heavy tail. The results indicate that the risk of extreme default realizations is more common than benchmark corporate default risk model estimation, raising concerns over the potential threat to international financial stability in terms of underestimating corporate default risk, especially during crisis periods.

Based on the results that global corporate default events are excessively clustered, I identify a novel source of default clustering: International corporate default contagion. I begin by using the Poisson Autoregressive model to assess the severity of corporate default contagion. Across all four regions, the Poisson Autoregressive model is separately estimated based on default data in time rescaled intervals. This approach provides a novel perspective for quantifying the impact of default contagion after controlling for the impact of firm-specific and systematic variables. The empirical results largely show that corporate default events exhibit significant results of persistent cross-border corporate default contagion at a regional scale. I subsequently extend the Poisson Autoregressive model to a multivariate setting. The empirical results confirm a considerable degree of cross-region default contagion, raising concerns of regional corporate vulnerability triggered by corporate default contagion from external regions.

Systemic risk drives the excess clustering of regional corporate default events rather than random intermittent surge of default events over time. Using Linton and Whang (2007) measure, or Han et al. (2016) a similar multivariate measure, I show that a high quantile right tail surge of default events elevates the risk of similar extreme waves of default events over multiple periods up to two years. The findings hold for all regions based on default data in time rescaled interval.

Finally, I identify a novel source that drives the excess clustering of global corporate default events: International corporate default contagion. I begin by using the Poisson Autoregressive model to assess the severity of corporate default contagion. Across all four regions, the Poisson Autoregressive model is separately estimated based on default data in time rescaled intervals. This approach provides a novel perspective for quantifying the impact of default contagion after controlling for the impact of firm-specific and systematic variables. The empirical results largely show that corporate default events exhibit significant results of persistent cross-border corporate default contagion at a regional scale. I subsequently extend the Poisson Autoregressive model to a multivariate setting. The empirical results confirm a considerable degree of cross-region default contagion, raising concerns of regional corporate vulnerability triggered by corporate default contagion from external regions.

Related Literature: This paper contributes to three strands of literature. First, I contribute to the strand in corporate default risk assessment based on reduced-form models. Research in this area largely focuses on identifying the relevant explanatory variables that impact default risk or are useful in predicting firms' distress risk. Altman (1968), Ohlson (1980), Zmijewski (1984) studies the relevant balance sheet variables that can be useful in predicting default risk. Shumway (2001), Chava and Jarrow (2004), and Campbell et al. (2008) point out that market-based variables can be useful in default risk assessment. Apart from firm fundamentals, Lando and Nielsen (2010), Duan et al. (2012), Koopman et al. (2011), and others identified a vast list of domestic macro-financial variables that impacts corporate distress risk. On top of Campbell et al. (2008) explanatory variables, Asis et al. (2021) point out that incorporating global financing variables can improve default risk assessment. This paper does not contribute to the corporate default risk literature by identifying new explanatory variables that improve default risk assessment. Instead, the paper shows that firm-specific and systematic variables, thus far identified in corporate default risk literature, are inadequate in accounting for the clustering of global corporate default events exhibited in data. The finding raises concerns that systemic risk poses a threat to corporate debt distress and advocates the importance of considering correlated default assessment methods in studying corporate default clustering, which relates to my contribution in the second strand of research.

Second, this paper contributes to the strand on the clustering of corporate default events. Following Das et al. (2007) pioneer findings that their four explanatory variables are inadequate in accounting for corporate default clustering in the U.S., default clustering literature has since relied on two main approaches to account for this phenomenon. The first approach is a time-varying latent variable (frailty) that drives the clustering of corporate default events. Duffie et al. (2009) points out that introducing a frailty factor provides a more realistic assessment of default risk. From the out-of-sample perspective, they show that considering a frailty factor provides a wider default risk forecast confidence interval. This approach allows them to potentially account for the extreme realization of default events during crisis periods, which cannot be obtained without incorporating the frailty factor. Koopman et al. (2011) also identifies a similar result but estimates the frailty factor using a different econometric model. The second approach relies on default contagion. Based on aggregated data at the economy level, Azizpour et al. (2018) use a point process model to show that incorporating both frailty and contagion are necessary to explain the total degree of corporate default clustering, as realized in the U.S. economy. This paper contributes to the default clustering literature by identifying a novel source of clustering: International corporate default contagion. I quantify the impact of international corporate default contagion on default risk after controlling for firm-specific and systematic variables. Across different regions worldwide, I obtain this objective by assessing the severity of default contagion based on default events in time rescaled intervals. Apart from identifying highly significant degree of cross-border default contagion at a regional scale, I also identify evidence of cross-region default contagion. Notably, my approach differs from Azizpour et al. (2018), which omitted the information in firm-specific variables.

Third, this paper contributes to the strand on financial risk contagion. Studies on financial contagion can be classified into corporate finance and international finance. In the corporate finance area, Lang and Stulz (1992) identifies evidence of bankruptcy contagion in the U.S. based on adverse movement in stock returns. Using stock returns and CDS spreads, Jorion and Zhang (2007), and in a later work Jorion and Zhang (2009) also identify a similar finding. From the perspective of financial firms, Helwege and Zhang (2016) points out that financial default contagion can occur from both the counterparty and informational channels. Instead of relying on market-based variables, Azizpour et al. (2018) identify significant evidence of default contagion among firms in the U.S economy, even after controlling for the relevant systematic variables and a frailty factor. Notably, Azizpour et al. (2018) relies on actual corporate default events rather than a proxy of firms' credit risk.

From the international finance perspective, a large body of research has identified significant empirical evidence of sovereign default risk contagion during financial crisis periods. For instance, based on European sovereign CDS data, Aït-Sahalia et al. (2014) and Blasques et al. (2016) identify significant evidence of sovereign default risk contagion during the 2008 Global Financial Crisis and Eurozone Sovereign Debt Crisis. Candelon and Tokpavi (2016) also identify a similar finding, but their analysis is based on European stock market indices data. Elevated sovereign default risk may increase default risk of firms operating in a financially distressed country. However, most existing studies in international finance exclusively focus on the macro aspect of financial risk contagion at the country level. Minimal studies explicitly focus on the severity of micro firm-level financial risk contagion across different borders. Closer to our research, Hassan et al. (2021) show empirical evidence that elevated perception of sovereign risk impacts multinational firm decision at the micro level. These include the reduction of investment and employment in the relevant country of concern. However, Hassan et al. (2021) did not directly focus on the impact of the contagion of firm-level financial risk (or default risk) across different economies.

This paper aims to contribute to the financial risk contagion literature from both corporate finance and international finance perspectives. Instead of focusing on corporate default clustering in a closed economy setting, I focus on evaluating the severity of default clustering across different regions on a global scale. In this process, I identify and quantify international corporate default contagion as a novel source of corporate default clustering in the U.S. economy, and the rest of the world.

The paper proceeds as follows. Section 2 justifies the rationale for studying corporate default clustering from an international perspective. Section 3 discusses the methodology for evaluating the severity of corporate default clustering. Section 4 presents the data and the main empirical results. Section 5 discusses the approach to assessing the severity of international corporate default contagion. Section 6 concludes.

2 International Corporate Default Contagion

This section justifies the economic rationale of studying corporate default clustering from an open economy macroeconomic setting. As mentioned in the previous section, prior literature on corporate default clustering approaches this topic from a closed economy perspective. Instead, I propose the surveyance of this topic from a global perspective by referring to a large body of research in corporate finance, international trade, and international finance. My approach allows for the identification of a novel source of corporate default clustering: International Corporate Default Contagion.

Azizpour et al. (2018) point out the complex and intricate interdependence among firms based on the supply chain and production network channel. They list several research to support this viewpoint, such as Boone and Ivanov (2012), Acemoglu et al. (2012), Elliott et al. (2014), to name a few. Based on these studies, Azizpour et al. (2018) justify that a default by a critical firm (or several firms) may impose a financial burden on other surviving firms, thus justifying for corporate default events to be contagious. However, the key assumption in their research lies in focusing on the impact of default contagion from a closed economy perspective. In reality, firms develop extensive business and financial relationships across borders, and even to some extent, on a global scale. Global value chain literature explicitly put forth this expansive relationship. Antras et al. (2017) point out that multinational firms may source input materials globally. In this case, default or failure by key firms along the global value chain may disrupt economic activity for the remaining surviving firms, elevating the financial distress risk of these firms. Other studies in global value chain, such as Antràs and Chor (2013), Antràs and De Gortari (2020) portray a similar intuition.

Using a structural model supported by empirical results, Chen et al. (2020) show that financially distressed firms have less incentive to be patient and compete aggressively over short-term gain. This intensified competition may reduce profit margins for all firms in the same industry, elevating financial distress risk for all firms in the same industry. By extending their theoretical model to a multi-industry setting, Chen et al. (2020) also shows that intensive competitive behavior by distressed firms operating in multiple industries can also elevate distress risk for surviving firms across these industries. Chen et al. (2020) findings have adverse implications for the global economy, where multinational firms compete internationally. In this setting, distressed firms may compete intensively. This aggressive competition creates negative externalities for surviving firms across different economies. In this case, firms face an increased risk of financial distress, even though they may operate in different economies. Using a network model, Dou et al. (2021) also identify a similar finding.

Apart from tangible business and financial linkages across firms, information contagion is another channel that triggers corporate default at a cross-border scale. Using a global games approach, Oh (2013) illustrates that default by a critical firm in the economy could reveal harmful information, which increases distress risk for other surviving firms. When a critical firm in the economy has defaulted, creditors believe that other non-related firms in the economy could also be exposed to the same risk and practice discriminatory lending practices against related surviving firms, even if these firms may have strong financial fundamentals. The discriminatory lending practice reduces credit supply for remaining surviving healthy firms, thus elevating the financial distress risk of these firms, facilitating coordinated waves of corporate default events. Based on a structural model of multi-firm default, Giesecke (2004) also portrays a similar intuition.

Related empirical studies also show that revealing negative news on a specific firm may adversely impact related firms. Lang and Stulz (1992) and Jorion and Zhang (2007) show that bankruptcy announcements negatively affect firms in the same industry. Schwenkler and Zheng (2020) propose a machine learning method to identify a similar finding based on a decline in stock price and increase in credit spread of firms, which may not be directly involved in the news report. The 1997 Asian Financial Crisis is one of the real-world examples illustrating the severity of information contagion. Since Thailand's currency market failed, international investors gradually became more pessimistic about the Asia Pacific region. Loss of confidence gradually spread across most economies in the Asia Pacific Region, triggering a massive wave of capital outflow from the region. The credit crunch gradually worsens, leading to a large wave of corporate default events in the region.

A recent credit crisis episode in China also raises concern about the severity of international financial contagion that arises from global firm interlinkages and the information contagion channel. Since negative reports of Evergrande's financial distress were revealed, global investors and financial institutions became largely adverse in providing funds to real estate firms in China. Firms in other related industries are also affected. Lenders were concerned that the firms in related industries may also hold hidden debt that may not be reflected in their balance sheets. Consequently, these firms face tremendous challenges in raising funds, as reflected in exorbitant costs in raising new debt or refinancing existing borrowings. As of writing, the debt crisis in mainland China's real estate industry has triggered an unprecedented wave of defaults by firms in the related industries and may potentially drive more firms into financial distress in the near future. This trend is a severe cause for concern, as some of these firms previously showed minimal symptoms of financial distress.[1]

[1] For instance, Fantasia Holdings defaulted on Oct 2021. Zhenro Properties Group Ltd raised concerns that it might not meet its financial obligations in Feb 2022, and defaulted in April 2022. Shimao Group Holding Ltd

In a financial stability report by the Federal Reserve on November 2021, the Fed suggested the potential of adverse global financial contagion due to Evergrande's financial fallout. As Evergrande is one of the largest developers in China, the report raised concerns that Evergrande's financial distress can impose severe stress on the Chinese financial system. Due to the size of China's economy and its extensive trade linkages in the global economy, Evergrande's financial distress can also potentially impose a tremendous strain on global financial markets. In this case, multinational firms worldwide that have direct and indirect business linkages with the real estate industry in China may be adversely affected. Consequently, potentially facilitating additional waves of corporate default events on a global scale. Former Fed Chair Janet Yellen also echoed a similar concern in November 2021.[2]

International finance literature also hints that 'contagious' corporate default events can be a threat to global financial stability. Specifically, firms' financial distress adversely impacts the country's default risk that it domiciled in. Sovereign default risk can be contagious across economies during a crisis period. Elevated sovereign default risk, in turn, affects domicile firms' borrowing costs and financial distress risk. However, the global default 'contagion' risk factor may not be adequately reflected in explanatory variables. In this case, excess clustering of corporate default events may occur due to omitting the international default contagion factor that drives firms' financial distress. Both Wu (2020) and Kwak (2020) provide empirical evidence that increase in corporate debt leads to the increase in sovereign default risk. Based on this insight, they independently construct a structural model to show that elevated corporate default risk reduces firms' ability to pay tax revenue, leading to an increase in sovereign default risk.

A large body of literature has documented empirical evidence of sovereign default risk contagion during a financial crisis. Aït-Sahalia et al. (2014) use a multivariate Hawkes model to show evidence of contagion among Euro Area sovereign CDS spread during the Eurozone Sovereign Debt Crisis period. In a similar period, Candelon and Tokpavi (2016) also show significant contagion effect in Europe by showing the common extreme decline in stock indices. Elevated sovereign default risk exposure increases domestic firms' risk of not fulfilling their financial obligations. Ağca and Celasun (2012) point out that a sovereign can transfer resources from the corporate sector to finance its fiscal needs during fiscal distress, suggesting that the fiscal outlook of a country determines firms' borrowing costs. Empirical studies in sovereign default risk also point out sovereign risk impact pricing of corporate bonds (e.g. Bedendo and Colla (2015); Bevilaqua et al. (2020), etc). Hassan et al. (2021) provide empirical evidence that elevated perception of sovereign risk leads to reduced investment and employment in the respective country of concern. Overall,

defaulted in Jul 2022. These default events are a cause for concern as the firms have a relatively healthy balance and previously showed minimal signs of entering into financial distress.

[2] Yellen stated that an economic slowdown in China would have "global consequences".

these firm-economy relationship suggests that default by firms in a specific country could adversely elevate the default risks of firms in other regions.

Additionally, firms in developed economies often have incentives to access capital markets by issuing securities through cross-border affiliates. This approach of gaining access to funding circumvents capital control, reducing taxes or other forms of payment. Coppola et al. (2021) propose a novel approach to restate firms' bilateral investment positions. Using this approach, they discover that investment from firms in developed economies to firms in emerging markets is significantly larger than previously documented. The finding suggests that corporate financing activity in emerging markets is considerably more affected by activities in the developed economies than previous thought, vice versa. In this case, a crisis in developed economies may adversely impact firms in emerging markets access to funds, potentially facilitating cross-border default contagion across different continents. Firm-specific and systematic variables may not adequately account for cross-border default contagion risks. In that case, both developed and emerging markets may display excess default clustering that explanatory variables could not explain.

3 Methodology: Econometrics Model and Time Rescaling Poisson Method

This section critically assesses for evidence of excess clustering in corporate default events that cannot be accounted for by firm-specific and systematic variables. I begin by introducing the benchmark corporate default intensity specification that measures individual firm's default risk exposure. Subsequently, I explain how the time rescaling method can be applied on our cumulative aggregate default intensity to assess for excess clustering in corporate default events. Suppose our default intensity specification is correctly specified. In that case, the count of default events in the new rescaled time interval should follow a Poisson distribution. Finally, I discuss a battery of econometrics tests for assessing the severity of excess default clustering.

3.1 Proportional Hazard Model

The benchmark econometrics model is based on a proportional hazard model. I measure the default intensity of firm i at time t, $\lambda_{i,t}$, as:

$$\lambda_{i,t} = \exp(X_{i,t}^T \beta) \tag{1}$$

In (1), β is the parameter estimated using the maximum likelihood estimation. X refers to the explanatory variables incorporated into the default intensity model. The explanatory variables include both firm-specific and systematic variables. In the next subsection, I critically discuss the relevant explanatory variables included in the proportional hazard model.

My default intensity specification (1) is similar to the benchmark econometrics model in corporate default clustering, such as Das et al. (2007), Duffie et al. (2009). Besides research in corporate default risk, a similar specification of the proportional hazard model is also used in other areas of research in risk management, such as operation risk assessment by Chernobai et al. (2011). However, this measure of default intensity is noticeably different from research in mortgage default risk, such as Deng et al. (2000), Clapp et al. (2006), among others. Mortgage default risk models tend to include a baseline hazard function $\lambda_0(\tau)$ intercept that varies with time τ. In my corporate default risk model, the time-varying baseline hazard function is excluded from the model as public corporations tend to exist perpetually, except for corporations going bankrupt or becoming delisted due to acquisitions or privatization. In contrast, mortgage loans are completely paid off after a finite time horizon. The inherent default risk exposure of individual mortgage loan naturally vary at different time period of mortgage loan, thus explaining the relevance of a time-varying baseline hazard function for assessing mortgage default risk.

3.2 Time Rescaling Poisson Theorem

This subsection explains the time rescaling method and the application of this approach in assessing if corporate default events display an excess degree of clustering. In the previous subsection, I propose a doubly stochastic approach to estimate corporate default risk exposure. This approach means that the corporate default intensity for firm i at time t, $\lambda_{i,t}$, is stochastic and solely determined by a range of firm-specific and systematic variables. Following Das et al. (2007), I explain how the time rescaling method can be applied in assessing if corporate default events display excess clustering beyond the information contained in explanatory variables.

By defining τ_i as the default time of firm i, I begin by modeling the cumulative number of default events up to time t, among a total of n firms as:

$$N_t = \sum_{i=1}^{n} \mathbf{1}_{\{\tau_i \leq t\}} \tag{2}$$

Correspondingly, the total default intensity of all surviving firms at time t can be written as:

$$\lambda_t = \sum_{i=1}^{n} \lambda_i(t) \mathbf{1}_{\{\tau_i \geq t\}} \tag{3}$$

Intuitively, equation (3) suggests that the aggregate default intensity for all surviving firms in the economy may be obtained by summing up the default intensity of each individual non-defaulted firms in the economy. Apart from the correlation among the explanatory variables that are incorporated in the default intensity model, the model is unable to account for any additional correlation among default events in the economy (and also across different economies). The total accumulative default intensity (compensator) over a specified time interval can be written as the integral of equation (3):

$$U(t) = \int_0^t \sum_{i=1}^{n} \lambda_i(s) \mathbf{1}_{\{\tau_i \geq s\}} ds \tag{4}$$

Based on the compensator, I can construct artificial time bins such that each time bin contains the same measurement of aggregated default intensity. To do so, I select an arbitrary number c, which is referred to as the bin size. I next select non-overlapping calendar time at time period t_0,..,t_K, where $t_0 = 0$, $t_K \leq$ T and

$$U(t_i) - U(t_{i-1}) = c \tag{5}$$

From equation (5), K refers to the total count of artificially constructed time bins based on the data across the entire sampling period. Correspondingly, I can then proceed with counting the number of corporate default events in the kth artificial time bin as:

$$X_k = \sum_{i=1}^{n} \mathbf{1}\{t_{k-1} \leq \tau_i < t_k\} \tag{6}$$

Based on Meyer (1971), accumulated default events in the constructed time bin will follow an independent Poisson process, with parameter c, if the relevant explanatory variables (firm-specific and systematic variables) are selected, and the corporate default intensity specification (1) is correctly specified.

To assess if the results are robust to the selection of different bin sizes, I select different counts of bin sizes from integers of 2 to 12, with a difference of 2 default counts at each interval. The objective is to avoid selecting a small bin size that constructs a time rescale interval with a short time interval. In this case, the arrival of corporate default events in each time interval is almost simultaneous, resulting in limited time for default events to accumulate. Thus, corresponding econometrics tests may not adequately assess the severity of default clustering. At the same, I also aim to avoid selecting a large bin size that results in a data sample with low observation. The latter case leads to practical issues in the application of econometrics tests on data in the rescaled time interval due to low data count.

After applying the time rescaling method on the aggregate default intensity measure, I conduct several econometrics tests to assess if corporate default events display an excess degree of clustering beyond the information in explanatory variables. The econometrics tests are largely based on Karlis and Xekalaki (2000), and also several econometrics tests that are proposed by Das et al. (2007), Lando and Nielsen (2010). Based on the aforementioned studies, I consider six different econometrics tests: (1) Fisher Dispersion Test, (2) PWB, (3) Autoregressive Test, (4) KK, (5) Upper Tail Test (Mean), (6) Upper Tail Test (Median). In this order, the econometrics tests enable us to study the dispersion, distribution, independence, generating function, and tail properties of the corporate default data in our time rescaled interval. Suppose corporate default events display an excess degree of clustering beyond the information in explanatory variables. In that case, default data in time rescaled data will exhibit severe degree of dispersion, have heavy tail, display a strong degree of serial dependence, and should not follow a Poisson distribution. In the Appendix section A.1, additional details on the background and purpose of each econometrics test are explained in further detail.

4 Data and Primary Empirical Results

4.1 The Data

My dataset contains a global coverage of firm-level data in different countries/economies across different continents worldwide. Specifically, the dataset includes information on corporate default events, as well as a comprehensive coverage of firm-specific and systematic variables.

The main source of my data is retrieved from the CRI database, the Credit Research Initiative at the National University of Singapore (NUS CRI), accessed on July 1, 2021. The NUS CRI database contains information on corporate default events, accounting, and market-based data for over 70,000 publicly listed firms in 133 countries/economies from 1990 onwards. However, data coverage for firms before 1995 is limited for most economies. As such, my analysis focuses on data from January 1996 to December 2020. To provide an adequate assessment on the severity of corporate default clustering on a global scale, my data sample covers 34 economies across different continents.[3] As corporate default events are rare, several countries/economies may not contain sufficient coverage of corporate default events. This may hinder us from deriving meaningful statistical insight if we intend to assess the severity of default clustering in a specific economy. To mitigate this issue, we group several countries/economies at a regional level based on geographical proximity or similarities in the structural characteristics of the economy. Overall, my research focus on corporate default risk analysis in four regions: U.S., Europe, Asia Pacific, and the Emerging Markets regions. Table 1 presents the economies included in each of these regions.

Apart from global coverage of corporate default events, another distinct aspect of our

[3] While NUS CRI database may contain data for a large number of economies worldwide, data in most of the economies are sparse. To ensure sufficient data for data analysis, we only consider economies with an average of over 100 firms each year and at least one default event over the entire data sample.

dataset is that it contains background information on each corporate default event. Specifically, for each corporate default event, I can classify them into three main categories: (1) Bankruptcy, (2) Default, (3) Debt Restructuring. Within these categories, I can further classify them into additional subcategories, which provide additional information on the nature of the specific corporate default event. Table 2 shows the breakdown of this classification. This information is useful for the analysis as different countries/economies have variations in bankruptcy laws and may differ in the definition of corporate default events. To be consistent in our classification of distress indicators across different economies, I do not classify bailout or delayed payments made within grace period as an indicator of financial distress.

In analyzing the severity of default clustering, one of the most crucial aspects lies in selecting the relevant explanatory variables. My benchmark explanatory variables are: Distance to Default, Three Month Rate (TMR), Trailing Stock Return, Trailing Stock Indices[4]. These are the explanatory variables used by Das et al. (2007) in their proportional hazard model. According to Das et al. (2007) and Duffie et al. (2009), these explanatory variables provide comparable out-of-sample predictive performance to existing documented research in corporate default risk assessment. However, the above variables strikingly omitted firm distress risk information that are available in firm-level accounting variables and other macroeconomic conditions. In this case, their approach inevitably omits useful information in assessing corporate default risk. For instance, they do not explicitly include variables that measure a firm's profitability or liquidity, which provides crucial information on a firm's default risk exposure. Consequently, solely relying on Das et al. (2007) four explanatory variables may lead to an overestimation of the severity of excess default clustering due to omitted variables. In other words, omitting critical explanatory variables may lead to the incorrect conclusion that the data suggest evidence of excessive default clustering, which may be corrected if we correctly specify our default intensity model. Lando and Nielsen (2010) addresses these shortcomings by showing that including additional explanatory variables results in less severe excess default clustering. However, their study solely focuses on the U.S. economy during a relatively benign time period (excluding global financial crisis). Moreover, they do not consider the severity of default clustering outside of the U.S. economy. Lando and Nielsen (2010) study raise additional questions over whether default events are still excessively clustered, if we focus on a different time period that includes financial firms, or focus on the severity of excess default clustering in other regions. In our study, we aim to address these concerns.

To mitigate this issue, I consider additional firm-specific variables that were proposed by Shumway (2001), Campbell et al. (2008). Specifically, I incorporate the market-to-book ratio, net income to market-value total asset, and cash to market-value total asset. The first metric measures a firm's market value, relative to its book value. The latter

[4] TMR and stock indices for the respective economies are based on NUS-CRI (2021)

two variables are a measure of a firms's profitability and liquidity. Our profitability and liquidity measure follows Campbell et al. (2008). I scale net income and cash by the market value of assets rather than the book value of asset. Campbell et al. (2008) justify this approach to be superior for being able to more efficiently incorporate the intangible aspects of the firm, as compared to the book value of the firm. I also consider additional global financing variables and global factors that were considered by Lando and Nielsen (2010), Azizpour et al. (2018), Asis et al. (2021). These include the yield curve, global GDP growth rate, and oil price. Table A.1 in Appendix provides additional details on the construction and source of our explanatory variables. To reiterate, the additional explanatory variables considered in the proportional hazard model allows us to account for firm-specific information and global macroeconomic condition that provides information on firm's exposure to distress risk. This information may not be accounted for by the economy's three month rate and market-based information that was considered by Das et al. (2007) and Duffie et al. (2009).

Unlike Das et al. (2007) and Lando and Nielsen (2010) a key distinction in my study lies in the inclusion of financial firms in our firm-level data sample. Existing studies in corporate default clustering based on firm-level data tend to exclude financial firms from their data sample. This approach is necessary due to distinct differences in capital structure among financial and non-financial firms. Specifically, financial firms tend to hold additional debt, such as deposit accounts, which are not classified under their current liability or long-term debt. As such, incorporating non-financial and financial firms in the same data sample raises concerns that empirical results may be distorted, leading to poorer default risk prediction.

However, financial firms play a critical role in the functioning of the overall economy. As such, excluding financial firms from the data sample may lead to an underestimation of the severity of corporate default clustering. Former Chairman of Federal Reserve Ben Bernanke has pointed "The failure of large, complex, and interconnected financial firms can disrupt the broader financial system and the overall economy, and such firms should be regulated with that fact in mind".[5] A large body of literature has also pointed out the systemic importance of financial firms in the overall economy, and the failure of financial firms may trigger financial instability. Acemoglu et al. (2015) point out a complex network relationship of financial firms in an economy, and a sufficiently severe shock to a critical financial firm (or firms) may give rise to financial instability in the overall economy. Bernanke et al. (1999) propose the financial accelerator channel, which also illustrates dependence in default risk among financial and non-financial firms. This mechanism points out that shocks that give rise to distress in financial firms may amplify and be transmitted to the real economy. In this case, defaults by key financial firms in the economy may generate adverse shocks, leading to an increase in default risk of non-financial firms.

[5] 2010, Squam Lake Conference, New York, June 16th

Compilation of corporate default events pertaining to financial firms also confirms the aforementioned financial instability concern. Figure 2 depicts the default events associated with financial firms in the four regions. Defaults by financial firms may be generally rare. Based on our data sample, in Developed Economies (U.S., Europe, and Asia Pacific), default rate of financial firms is about 0.03%. In Emerging Economies, the default rate of financial firms is slightly higher, at about 0.1%. However, Figure 2 also shows that most of the financial firms' default events tend to cluster during an adverse financial period. Specifically, across all four regions, the bulk of financial firms' default events cluster during the early 2000s and global financial crisis. At a regional level, most of the financial firms in Europe defaulted over the span from 2012 - 2013, coinciding with the peak of the Eurozone Debt Crisis. In Asia Pacific and Emerging Markets region, the data shows a considerable number of default events associated with financial firms during the late 1990s. This period coincides with the Asian Financial Crisis of 1997, as well as the Russian Default and Long-Term Capital Management (LTCM) crisis. Overall, from a global perspective, default by financial firms raise concerns that omitting financial firms from our data sample will lead to underestimating the severity of default clustering.

I can include financial firms in our data sample by using a non-standard computation of the Distance to Default (DtD) measure based on Duan et al. (2012). To justifiably combine financial and non-financial firms in the same data sample, Duan et al. (2012) propose a method for computing DtD by including an additional unknown parameter to adjust for additional debts of financial firms in the DtD measure. This unknown parameter accounts for the intrinsic differences in liabilities between financial firms and non-financial firms, which can be estimated using the maximum likelihood estimation. Using this approach for computing the DtD, Duan et al. (2012) shows that their approach provides commendable default risk predictive performance for financial firms. This result allows them to justifiably combine financial and non-financial firms within the same data sample, with minimal compromise in the overall predictive performance of corporate default risk.[6]

Overall, the firm-specific variables that I select are: Net Income to Market Value of Total Assets (NIMTA), Cash to Market Value of Total Assets (CASHMTA), Market to Book Ratio (MB), Distance to Default (DtD), and Stock Return (Trail). The systematic variables that I select are: Three Month Rate (TMR), Stock Index Return (Trail), Yield Slope, Oil Price, and Global Growth Rate. In Table A.1 of the Internet Appendix, I provide detailed information behind the source and construction of the explanatory variables.

Firm-specific variables are winsorized at the 1st and 99th percentile of their distribution.[7] This step aims to control for potential errors and eliminate unusual outliers in the

[6]See also NUS CRI Technical Report (2021) for application of a similar approach on an international data sample

[7]Market-to-book ratio is winsorized at the 5th and 95th percentiles due to concerns over numerous instances of small or negative book-to-equity values, which in turn make MB very large. Asis et al. (2021) also conduct a similar approach.

balance sheet and market data. Accounting Ratios (NIMTA, CASHMTA, and MB) are lagged by two months to ensure that accounting information is available at the time of predicting default risk. Systematic variables are not lagged or winsorized.

I present the summary statistics of the explanatory variables for all four regions. Table 4 presents the summary statistics for the firm-specific variables. Similarly, Table 5 presents the summary statistics for the systematic variables. In both tables, I provide the summary statistics for our full sample and a subset sample that only include defaulted firms. The indicator for defaulted firms is measured in the month before the default event (t-1). Our summary statistics include mean and a t-test analysis which evaluates if there is a statistically significant difference in means between the full sample and defaulted firms. Our summary statistics is computed based on the equal weightage of firms across time periods.

Among all four regions, Table 4 shows that defaulted firms generally face the following conditions: less profitable (NIMTA), less capable in covering short-term financial obligations (CASHMTA), higher market to book ratio (MB), lower distance to default (DtD), and negative trailing stock return. These summary statistics are largely consistent with economic intuition, as well as the results documented in Campbell et al. (2008), Asis et al. (2021), Lee (2022), etc.

4.2 Likelihood Estimation

Table 6 presents the parameter (β) estimates for our explanatory variables and the statistical significance of the parameter estimates. Our results cover the four regions: U.S., Europe, Asia Pacific, and Emerging Markets. Even after incorporating the new firm-specific and systematic variables, the four key explanatory variables considered by Das et al. (2007), except for economy stock indices, are largely significant across the four regions. From the perspective of the U.S. economy, this finding is consistent with the result documented by Duffie et al. (2009).

The parameter estimates of the newly introduced firm-specific and systematic variables are largely significant across the four regions, mainly at the 1% level, and also show the appropriate sign. This observation justifies the inclusion of our proposed explanatory variables in assessing the severity of corporate default clustering. For instance, the parameters for my measure of profitability and liquidity are mostly negatively significant. This observation highlights that an increase in profitability and liquidity reduce the risk of a firm falling into financial distress. Market-to-Book ratio is positively significant for most regions, suggesting that an increase in Market-to-Book tends to elevate a firm's distress risk. Firms in financial distress tend to face losses resulting in lower book value of equity, thus driving up the Market-to-Book ratio. These results are also consistent with our findings in Table 6.

Three Month Rate is negative and significant for all regions, except for the Asia Pacific region. An increase in Three Month Rate may increase a firm's borrowing cost, which elevates a firm's risk of financial distress. However, Central Banks worldwide tend to reduce interest rates during a macroeconomic recession, justifying the positive relationship between Three Month Rate and a firm's distress risk. We observe that the U.S. yield slope, a global financing variable, is positive and significant for all regions except for the U.S. economy. An increase in yield slope leads to a risk-off environment and raises firms' borrowing costs worldwide. Moreover, if firms hold a high proportion of debt denominated in U.S. dollar, an increase in U.S. interest rate may expose firms to adverse currency movements, increasing the risk of the firm's financial distress. However, an opposite but significant sign is observed in the U.S. economy. This may be the case where U.S. central bank tends to increase the interest rate if the economy is performing well.

Oil price is significant and negative for most regions at the 1% level. This is because an increase in oil price generally signals that the global economy is performing well. It may also indicate rising inflation, reducing firms' risk of defaulting on their financial obligations (e.g. Bhamra et al. (2010), etc). The coefficient estimate for oil price is negative and significant for all regions, except for Europe. An increase in the global economic growth rate creates a favorable international business environment and reduces firms' risk of falling into financial distress. This variable is negative and significant in the U.S. economy but not for Europe and Asia Pacific. However, the coefficient estimate for global economic growth rate is positive and signficant in the Emerging Markets. As such, this variable is excluded from the Emerging Market region. While I may use a different econometric model, my overall findings are also largely consistent with the results that are documented by Campbell et al. (2008), Asis et al. (2021), and Lee (2022).[8]

To further justify that the inclusion of our additional firm-specific and systematic variables is appropriate, we also conducted an out-of-sample (OOS) analysis. In our OOS analysis, parameters of the proportional hazard model are updated at a monthly interval recursively. I first estimate the parameters of the default risk models, using our data sample from Jan 1996 – Dec 2006. Subsequently, I start to predict default risk in the following month in a recursive approach. The Area Under the Curve (AUC) for the U.S. economy, Europe, Asia Pacific, and Emerging Markets are 0.968, 0.921, 0.929, and 0.827. These results show a noticeable improvement, compared to the OOS conducted based on Das et al. (2007) four explanatory variables. In table A.2 of the Appendix section, I compare

[8]Similar to Asis et al. (2021), I am also concerned about multicollinearity issues among market-based variables and global financing variables. To assess for multicollinearity issues, I conduct the Variance Inflation Factor (VIF) analysis among explanatory variables. VIF is calculated as $\frac{1}{R_j^2}$, where R_j^2 is the R^2 value of factor j on the rest of the explanatory variables. VIF values above 10 suggest evidence of multicollinearity concerns. In our default risk models, across all regions, we observe that our VIF across explanatory variables in different regions is largely below 5. This result suggests no serious multicollinearity issue among our selection of explanatory variables. As such, we did not calculate the market-based variables based on the residuals approach that was used by Asis et al. (2021).

the OOS performance based on our augmented model and the proportional hazard model that is based on Das et al. (2007) four explanatory variables.

As discussed earlier, the main focus of this paper lies in evaluating if a comprehensive selection of explanatory variables can adequately account for the clustering of corporate default events. To provide a preliminary assessment of the severity of default clustering across different regions, I provide a plot that compares the aggregate default intensity and actual corporate default events. Moreover, I also compare the aggregate default intensity computed based on our explanatory variables with Das et al. (2007) explanatory variables. Figure 3 presents the plot. Under a benign period, computed aggregate default intensity may not have displayed an obvious difference compared to the estimated aggregate default intensity based on Das et al. (2007) four explanatory variables. However, Figure 3 shows a noticeable improvement in the aggregate default intensity estimation under adverse macroeconomic environments, such as the early 2000s period. As a complement to OOS analysis, this observation further justifies the suitability of adding additional explanatory variables. Nonetheless, across all four regions, Figure 3 still shows that the aggregate default intensity estimate generally underestimates default risk exposure under a dire macroeconomic environment. At a regional level, for instance, the Europe region displays excess clustering of default events from 2012 to the mid-2010s, coinciding with part of the Eurozone Sovereign Debt Crisis period. In the Asia Pacific region, a similar phenomenon is observed during the late 1990s, coinciding with the Asian Financial Crisis period.

4.3 Econometrics Test for Excess Default Clustering

The findings in the previous subsection, particularly Figure 3, indicate several instances where corporate default events display an excess degree of clustering. Specifically, a rigorous selection of explanatory variables cannot adequately account for the clustering of default events. This observation holds for all four regions and tends to be more severe under crisis period. In this subsection, I present the empirical results of our econometrics tests on all four regions that aim to assess if corporate default events display degree of excess clustering. To reiterate, corporate default data is organized into time rescaled intervals based on subsection 3.2.

Table 7 and 8 compare the theoretical and empirical moments of corporate default data under time rescaled intervals across all four regions. Table 7 presents data for the U.S. and Europe regions. Table 8 presents data for the Asia Pacific and Emerging Market region. These tables also present the distribution of default events across several different bin sizes for each region. Across all bin sizes and regions, Table 7 and 8 show that the empirical mean of our estimation is largely close to the theoretical mean. This observation confirms that the proportional hazard models are reasonably well estimated in terms of aggregate corporate default risk estimation. However, across all four regions, data shows that empirical variance is noticeably larger than the theoretical variance of the Poisson distribution, especially at larger bin sizes. These results raise concerns that actual extreme realizations of corporate default events can be substantially larger than the default risk model estimation. The findings generally hold for all four regions, raising financial instability concerns due to underestimating actual default risk exposure in a region.

To complement the empirical results in Table 7 and 8, Figure 4 provides a plot that compares the distribution of corporate default data in time rescaled intervals and its theoretical Poisson distribution counterpart. Figure 4 shows the distribution of corporate default data in the U.S. and Emerging Markets based on bin size 6. Correspondingly, default data in Europe and the Asia Pacific are based on bin size 4. Each region contains about $200 - 300$ data observations, corresponding to an average of $1 - 1.5$ in each time rescaled interval.

Figure 4 shows that corporate default data in all four regions exhibit severe evidence of heavy tail beyond the estimation of a theoretical Poisson distribution. This finding is severely more pronounced for Europe and Emerging Market regions. Compared to other regions, the severity of heavy tail is less pronounced for the Asia Pacific region but still provides some evidence of heavy tail. The heavy tail observation in all four regions provides evidence of excess clustering in corporate default events, which explanatory variables cannot adequately explain. As a robustness check, Figure 5 shows a plot of corporate default data based on bin size 8 in the U.S. and Emerging Markets and bin size 6 for Europe and Asia Pacific. Figure 5 shows even more severe evidence of heavy tail across all four regions. This finding is expected as larger bin sizes have a longer time to accumulate larger counts of default events attributing to the impact of omitted variables in default risk estimation.

Based on the plots in Figure 4 and 5, I next focus on evaluating the severity of excess default clustering by conducting several econometrics tests on corporate default data in time rescaled intervals. Table 9 provides a summary of the Econometrics tests. For brevity, I only include the tests' p-value and statistical significance and omit the test statistics. The Kocherlakota-Kocherlakota test evaluates the generating function of the data. Across all four regions and bin sizes, I document that the p-value of our data is mostly less than 1%. These results indicate that the generating function of our time rescaled default data does not fit the Poisson probability generating function. Potthoff-Whittinghill-Bohning (PWB) test assesses the assumption that the mean equals the variance. Except for a few instances in bin sizes with low default counts, the PWB tests generally reject the null hypothesis for all four regions. These findings indicate that the mean of the data is not equal to the variance.

The Fisher Dispersion test and the two upper tail tests evaluate the dispersion and tail properties of the data in time rescaled intervals. Similar to the PWB test, I observe that the former three tests largely reject the null hypothesis, especially at larger bin sizes from 6 default count onwards. These findings confirm that default data exhibits a high degree of dispersion and heavy tail properties, thus not following a Poisson distribution.

As a complement to the results in Table 9, Table 10 presents the parameter estimates of the AR(1) model across different regions. As described in further detail in Appendix A.1, The AR(1) model can be written as:

$$X_t = A + BX_{t-1} + \epsilon_t$$

where A and B are parameters to be estimated, based on i.i.d demeaned Poisson random variables ϵ_t. To reiterate, the AR(1) model aims to assess if default events in time rescaled interval display evidence of serial dependence (B = 0). Across different regions and bin sizes, we reject the null hypothesis that the model intercept is equal to the bin sizes. Moreover, we also reject the null hypothesis that B = 0. This finding indicates that default data in time rescaled interval exhibits serial dependence, thus violating the assumption of independence in Poisson distribution. Test statistics of AR(1) model intercept (A) is computed by testing if the model intercept is equal to the bin size (c). Across all bin sizes and regions, the test rejects the null hypothesis that the model intercept is equal to the bin size. This finding further confirms that the data do not follow a Poisson distribution.

The empirical results in Table 9 and Table 10 show that Europe and Emerging Markets display the most severe form of default clustering. In these two regions, I observe that almost all econometrics tests are highly significant at the 1% level across all bin sizes. These results indicate that default events in the time rescaled interval do not follow a Poisson distribution, exhibits heavy tail, and display serial dependence. This result is expected due to the severe episodes of excess default clustering during part of the Eurozone Sovereign Debt Crisis period (2011 -2015), which is largely confined to Europe. Specifically, aggregate default intensity in Europe seems unable to account for the spikes in default events during this period. For the case of the Emerging Markets, our aggregate default intensity is unable to account for spikes of corporate default events during the financial crisis period in the late 1990s, and the global macroeconomic slowdown in the early 2000s.

Notably, in the U.S. and the Asia Pacific region, there are several instances where the econometrics tests do not reject the null hypothesis that the data follow a Poisson distribution. These results are reflected for low bin sizes. Nonetheless, they may not indicate that the proportional hazard models in these two regions are correctly estimated. By fixing aggregate default intensity at small bin size, leading to the construction of a narrow time interval. In some cases, the time interval can be less than a month. In this narrow time interval, default events may not be given sufficient time to accumulate. As such, the distributional properties of data in narrow time intervals may not be reliable.

Similar issues are also documented in Lando and Nielsen (2010), Giesecke and Kim (2011), and Azizpour et al. (2018).

Nonetheless, the overall econometrics tests indicate that corporate default data in time rescaled interval exhibits heavy tail, especially at higher bin sizes, and do not follow a Poisson distribution. These results hold for all four regions. Our findings suggest a strong likelihood of encountering extreme waves of corporate default events that far outweigh our default risk estimation, especially during crisis periods. This finding is consistent with the plot of aggregate default intensity in Figure 3. These results raise concerns that systemic risk posed a severe threat to global financial stability in terms of underestimating global corporate default risk during a crisis period. This finding warrants further analysis into the sources of global corporate default clustering.

5 Systemic Risk

The recent and ongoing real estate debt crisis in China raised concerns that multiple waves of coordinated widespread financial distress among corporations may be a systemic problem. For instance, default by systemically important firms may trigger waves of corporate distress due to firm interlinkages and other firm network effects.

To reiterate, the econometrics tests in the previous section show strong evidence of heavy tail and serial correlation in the time rescaled default data for all regions. The results indicate that corporate default events display a severe degree of excess clustering beyond the information in explanatory variables. However, the tests, specifically the serial correlation test, only provide a linear measure of the dependency among default events over time. It may not explicitly inform if an extreme wave of firms' financial distress increases the risk of similar waves of coordinated firm failures over multiple periods or if these events are random over time.

This section investigates if systemic risk triggers the excess clustering of corporate default events in each region or if they are random over time. To measure systemic risk, I assess whether an extreme quantile wave of firms' financial distress may elevate the risk of additional waves of corporate distress over multiple periods.

5.1 Extreme default events serial dependence

In investigating whether systemic risk is a driver of corporate default clustering, I first assess for evidence of serial dependence among periods in time rescaled interval that registers an extreme quantile waves of corporate default events. To do so, I construct a binary sequence taking a value of 1 if the realized count of default events in each time bin (X) exceeds the 1- α quantile of time bin with parameter c, and 0 otherwise. To capture the extreme quantile wave of default events, α takes a small value, such as 5%, 10%, or 15%. The relationship is expressed as:

$$I_t(\alpha) = \begin{cases} 1 & X_t \geq q_\alpha(c) \\ 0 & Otherwise \end{cases}$$

Following Christoffersen and Pelletier (2004), I define the duration between two consecutive extreme waves of default events (takes a value of 1) as:

$$d_i = t_i - t_{i-1}$$

As discussed in the previous section, default counts across time bins follow a Poisson distribution and are identical and independent. In this case, d_i follows a geometric distribution. Following Candelon et al. (2011), I write the orthonormal polynomial for geometric distribution d and 'success' probability α recursively as:

$$M_{j+1}(d;\alpha) = \frac{(1-\alpha)(2j+1)+\alpha(j-d+1)}{(j+1)\sqrt{1-\alpha}} M_j(d;\alpha) - \frac{j}{j-1} M_j(d;\alpha)$$

where $M_1(d;\alpha) = 1$ and $M_0(d;\alpha) = 0$

Candelon et al. (2011) show that the above test statistics converge to χ^2 distribution with degree of freedom 1. The null hypothesis of the above tests suggests that tail events are independent, suggesting that tail events within the data sample are random. However, if the time gap between most realized tail events in time rescaled interval is small, the null hypothesis that tail events are independent will be rejected. In this case, tail default events exhibit serial correlation.

Using the above test statistics, I assess if binary tail events are serially dependent. Table 12 presents the results for different time bins across all four regions. Panel (a), (b), (c) report the value such that α takes value of 5%, 10%, or 15%. Across multiple bin sizes and regions, Table 12 shows that the null hypothesis that tail default events in a region is independent is rejected. The results confirm that an extreme wave of corporate default events increases the risk of similar adverse waves in the near future.

5.2 Quantilogram: Measuring Degree of Serial Quantile Dependence

In the previous subsection, empirical results inform that extreme waves of corporate default display a substantial degree of serial quantile dependence across multiple periods. Based on this finding, I aim to quantify the magnitude of serial dependence among extreme quantile default events in a region over multiple periods.

To do so, I use a modified variant of the quantilogram method proposed by Linton and Whang (2007). This method is a special case of Han et al. (2016) cross-quantilogram method that focuses on a single time series variable. Define a stationary time series variable as $\{x_{i,t},\, t \in \mathbb{Z}\}$, where i $\in$ {1,2,3,4}. $x_{i,t}$ relates to the count of corporate default events in region i, at time rescaled period t. Following, define the quantile for $x_{i,t}$ as $q_{i,t}(\tau_i) = inf\{v : F_{x_i}(v) \geq \tau_i\}$, where $\tau_i \in$ (0,1). The objective is to measure the degree of serial dependence for two separate extreme quantile tail events in the same region i, across multiple time lags, which may be represented as: $\{x_{i,t} \geq q_{i,t}(\tau_1)\}$ and $\{x_{i,t-k} \geq q_{i,t-k}(\tau_1)\}$, where k is an integer that measures number of lags.[9] τ_i is arbitrarily selected to account for varying severity of extreme waves of corporate default events in a time rescaled interval of a specific region. Following Linton and Whang (2007) and Han et al. (2016), I classify a tail event: $\{1[x_{i,t} \geq q_{i,t}(.)]\}$, as positive, which they dubbed as quantile hit or quantile exceedance process. In other words, 1[.] is an indicator function that takes a value of 1, if the count of default events in a particular time rescaled interval exceeds a specified value determined based on an arbitrarily selected quantile, zero otherwise. Following a modified variant of Han et al. (2016), the quantilogram measures the serial correlation between the quantile-hit processes across different time lags as:

$$p_\tau(k) = \frac{E[\psi_{\tau_i}(x_{i,t} - q_{i,t}(\tau_i))\psi_{\tau_i}(x_{i,t-k} - q_{i,t-k}(\tau_i))]}{E[\psi^2_{\tau_i}(x_{i,t} - q_{i,t}(\tau_i)]E[\psi^2_{\tau_i}(x_{i,t-k} - q_{i,t-k}(\tau_1)]} \tag{7}$$

for k $\in \mathbb{Z}_{\neq 0}$, and $\psi_\alpha(u) = 1[\text{u} > 0]$ - (1-α). The stated quantilogram in (7) quantifies the magnitude of serial dependence among two extreme quantile waves of default events across different time lags at a specified quantile level.

I use the quantilogram $\hat{p_{tau}}(k)$ to measure the serial dependence of tail events up to the lag of time period k = 12. As the quantilogram method measure dependence among tail events, I select bin size in each region that provides sufficient data sample for the subsequent econometric analysis. I begin by constructing bin size 4 for the U.S. and Emerging Markets region. Correspondingly, bin size 3 for the Europe and Asia Pacific region. Based on these bin sizes, I obtain a data sample that contains almost 200 to 300 data observations in our time rescaled interval. The objective is to balance having a reasonable number of data observations to apply the quantilogram method and avoid constructing over narrow time rescaled interval (such as < 1 month).

I consider three different values of τ_i for constructing the extreme quantile interval: τ_i = 0.8, τ_i = 0.85, and τ_i = 0.9.[10] Notably, setting τ_i = 0.8 and 0.85, on average, coincides

[9] It is useful to note that Linton and Whang (2007) and Han et al. (2016) definition of tail events and original construction of the quantilogram and cross-quantilogram in their paper differ from us. In their paper, they define a tail event as $\{x_{1,t} \leq q_{1,t}(\tau_1)\}$. In this approach, an adverse event occurs when the economic or financial variable falls below a specified quantile value. In contrast, this paper deal with an alternate scenario where an adverse event occurs if realized default events exceed a specified quantile. In a working paper version of Han et al. (2016) paper, they also studied a similar approach of constructing the cross-quantilogram that this paper adopts.

[10] I also consider a separate case τ_i = 0.95, and τ_i = 0.99. However, I observe less evidence of serial dependence among tail events in these alternate cases. This observation is expected due to limited data on binary events at the extreme quantile level.

with the 90th and 95th quantile of the theoretical Poisson distribution. This result is due to the heavy tail nature of the corporate default data in the time rescaled interval, documented in the earlier section. Based on Han et al. (2016), I also calculate the 90% bootstrap confidence intervals for no quantile dependence, using 1000 bootstrap replicates, to measure if the quantilogram is significant.

Figure 9 presents the quantilogram $\hat{p_{tau}}(k)$ of extreme tail events at $\tau = 0.85$ for the four regions: U.S., Europe, Asia Pacific, and Emerging Markets. Across these four regions, Figure 9 shows that corporate default events exhibit a severe degree of quantile serial dependence. Specifically, Figure 9 suggests that an extreme wave of corporate default events may adversely increase the risk of similar extreme waves of corporate default events in the subsequent periods. For instance, from the perspective of the U.S. economy and at $\tau = 0.85$, the quantilogram shows that an extreme surge of default events increase the risk of another similar wave of default events over the next few time rescaled intervals (on average 1.5 months) at a probability of above 0.4. This impact gradually decays over subsequent periods into the near future but remains significant at the bootstrap 90% confidence interval. A similar result can also be interpreted from the perspective of Europe, Asia Pacific, and Emerging Markets region.

As such, Figure 9 suggests that each extreme wave of default events in U.S., Europe, and Emerging Market elevates the risk of another wave of an extreme wave in the next period by almost 0.4. This impact remains largely persistent before decaying to zero after multiple periods. An extreme wave of default events in Asia Pacific also increases the risk of another similar wave of default events in the next period. However, it is less persistent than in the other regions.

Figure 10 depicts a similar plot based on different bin sizes of 6 for the U.S. and Emerging Markets region and bin size 4 for the Europe and Asia Pacific region. Both Figure 9 and Figure 10 show largely similar results. Serial dependence among extreme quantile events is highly significant for all four regions. Additionally, serial dependence among quantile events is generally more severe at the earlier lags but gradually declines over time.

6 International Corporate Default Contagion

Empirical results in the last section shows that global corporate default events display excess clustering beyond information contained in explanatory variables. In this section, I investigate if international corporate default contagion triggers the clustering of global corporate default event. To begin, I provide an autocorrelation plot of corporate default events in the time rescaled interval for all four regions. I specifically choose bin size 5 for the U.S. and Emerging Markets. For the Europe and Asia Pacific region, I choose bin size

3. These bin sizes are selected as they contain sufficient data for additonal econometrics analysis, which will be conducted later in this section. At the same time, the data also display a severe degree of excess clustering.[11] Figure 6 presents the plot. Across all four regions, autocorrelations are all significant and large at the initial lags. However, the autocorrelation gradually decays closer to zero as the lags increase. These findings indicate that corporate default events exhibit a severe degree of correlation over multiple lag intervals. Moreover, the results confirm that default data in time rescaled interval violates the Poisson distribution and also support the earlier findings that default data are serially dependent. Overall, these findings provide preliminary evidence that cross-border corporate default contagion is one of the sources of regional default clustering and justify further investigation into this area.[12]

6.1 Regional Corporate Default Contagion

To assess the severity of international corporate default contagion, I evaluate the severity of cross-border corporate default contagion at the regional level. To do so, I estimate a Poisson Autoregression (PAR) model based on corporate default data in time rescaled interval for each region. The PAR model is based on Fokianos et al. (2009), and can be written as:

$$\lambda_t = \omega + \sum_{i=1}^{n} \alpha_i Y_{t-i} + \sum_{i=1}^{n} \beta_i \lambda_{t-i} \tag{8}$$

In equation (8), λ_t depicts the expected number of default events at time t. Similarly, λ_{t-i} represents the expected number of default events that was previously computed at period t - i. Intuitively, this term also relates to the aggregate default risk estimation in a region. Y_{t-i} depicts the count of the number of default events in time t - i. ω represents the intercept term.

α_i accounts for the sensitivity of default risk estimation at period t to actual default events realized at a time of i lags ago. This term accounts for the potential contagion effect by allowing default events accumulated at i periods ago to adversely elevate default risk exposure (λ_t) at period t. β_i accounts for the receptiveness of aggregate corporate default risk estimation (λ_t) at period t to default risk estimation (λ_{t-i}) at a lag of i periods ago.

[11] Based on our earlier econometrics tests, corporate default data in this bin size display excess degree of clustering, and contain sufficient data for our further econometrics analysis. For brevity, I only present the plots for these bin sizes. Autocorrelation plots at other bin sizes display similar results. Except for Europe, which contains 198 observations, the rest of the regions contain about 200 to 250 observations points, which allows us to collect sufficient data for subsequent econometrics analysis.

[12] An autocorrelation plot for all four regions, based on bin sizes of larger numerical value, also provides similar findings. However, this approach gives rise to a collection of time series variables containing low data counts, which hinders us from collecting sufficient data for additional econometrics analysis. As such, we only report the results for bin size 5 for the U.S. and Emerging Markets and bin size 3 for the Europe and Asia Pacific region.

Implicitly, this term captures the impact of default contagion by allowing estimated default risk based on multiple periods ago to dynamically increase default risk exposure at period t. In other words, this term captures the persistence of corporate default events. If there is no significant evidence of corporate default contagion, both α_i and β_i will be insignificant. In the special case where our proportional hazard model is correctly estimated, there will be no evidence supporting excess clustering of default events. In this case, there will also be no evidence that supports default contagion. In other words, only the ω term should be significant at parameter c, where c corresponds to bin size c.

Without loss of generality and to allow our empirical results to be conveniently compared across different regions and bin sizes, I focus and only present the results for i = 1. Minimal material changes are observed if I extend our results to consider the cases where i is greater than 1. As stated in Fokianos et al. (2009) and Agosto et al. (2016), $\alpha + \beta < 1$, in order for the model to be stationary. In this case, if both α and β are significant, this result suggests that default events are contagious.

Overall, by estimating the parameters of the PAR model on corporate default data in time rescaled intervals, my approach is novel for dynamically quantifying the effect of corporate default contagion over multiple time periods from a global perspective. My approach of modelling default contagion have some similarities with Giesecke and Kim (2011), Azizpour et al. (2018), especially Agosto et al. (2016). These studies could also dynamically quantify the impact of corporate default contagion over multiple periods. However, they solely focus on the U.S.

Specifically, the economic intuition for quantifying the impact of contagion on default clustering is similar to the modified Hawkes model that was considered by Giesecke and Kim (2011), and Azizpour et al. (2018).[13] The latter two studies evaluate for the severity of default contagion by allowing default events that occur at a time period to increase the firm's risk of default in the next period. This contagious impact decays over time. Likewise, our PAR model also allows default events at an arbitrary time period to dynamically impact default risk over multiple time periods. Moreover, in the PAR model, the default contagion impact also decays over time at a rate of β, as shown in Equation (8).

Similar to my approach, Agosto et al. (2016) also use a similar variant of Poisson Autoregressive model proposed by Fokianos et al. (2009). However, there is a key distinction between my application of the Poisson Autoregressive model and the contagion model used by Giesecke and Kim (2011), Azizpour et al. (2018), and Agosto et al. (2016). I quantify the impact of default contagion based on default data in time rescaled intervals.[14] In con-

[13] Unlike Giesecke and Kim (2011), and Azizpour et al. (2018), we do not explicitly address the severity of default contagion based on defaulted debt. This impact has been partially accounted for by default events in time rescaled data estimated using firm-specific and systematic variables.

[14] As discussed in the previous section, default intensity for each firm is estimated based on the proportional hazard model (1), which controls for a wide array of firm-specific and systematic variables. Time rescaling method is subsequently applied on cumulated default intensity, which accounts for the information contained in

trast, the former studies assess the impact of default contagion based on aggregated default data in calendar time. My approach allows us to control the information contained in firm fundamentals and systematic variables while assessing the severity of default contagion on the remaining excess default clustering. If information on corporate default contagion is already incorporated into market-based or other firm-specific variables, then omitting firm-specific variables will overestimate the severity of default clustering. In other words, Agosto et al. (2016), Giesecke and Kim (2011), and Azizpour et al. (2018) quantification of the impact of default contagion, based on aggregated data on calendar time, may provide a distorted assessment on the severity of corporate default contagion.

6.1.1 Parameter Estimates of PAR model

Table 11 presents the parameter estimates for the PAR model. Based on corporate default data in time rescaled interval, which controls for the effect of firm-specific and systematic variables, our results strongly support empirical evidence of corporate default contagion across the four regions. Specifically, we observe that ω is largely insignificant across different regions and bin sizes. Moreover, parameter α and β are also largely positive and significant. These results hold for the four regions, and across most different bin sizes. To reiterate, these observations jointly suggest that corporate default events in time rescaled intervals are highly persistent and violate the assumption of a Poisson distribution.

From the perspective of the U.S economy and at bin size 6, which corresponds to about two months of data in each unit of time interval, an increase of a unit default events at period t will lead to an increase of default events by 0.4 unit in the next period. This adverse impact also dynamically elevates default risk in the U.S. for several subsequent periods, though in a more muted way, which decays at about 0.4 over each period. A similar result and intuition can also be derived for the impact of default contagion in the Europe and Emerging Markets perspective.

A similar result is also observed in Europe. To reiterate, I show that an increase in default event in period t severely elevates default risk in the subsequent period and several periods into the future.This impact gradually decays over time. More specifically, at bin size 6, an increase in a unit default event leads to an increase in default event by almost 0.5 in the next period. This impact will persistently elevate default risk over multiple periods, though it decays at about 0.3 in each period. An analysis of corporate default contagion in Europe based on different bin sizes also provides a similar finding. My finding is also consistent and complements a large body of empirical research that identify significant evidence of sovereign risk contagion across different economies in Europe economies (e.g. Blasques et al. (2016), Candelon and Tokpavi (2016), Aït-Sahalia et al. (2014)).

firm-specific variables.

Notably, my empirical results also support significant evidence of default clustering in the Asia Pacific region. However, evidence for persistent default contagion is less severe in this region. Across most of the bin sizes, this finding can be inferred from positive and significant parameter A, but not for parameter B. This result suggests that only default events in the last period impact corporate default risk in the Asia Pacific region for this specific period but not subsequent periods. In other words, default events are not persistent in Asia. This finding could be due to economies in the Asia Pacific region being relatively less integrated as compared to the U.S. economy or Europe.

6.1.2 Impulse Response of Corporate Default Events

As a complement to the parameter estimates of the PAR model, it may also be useful to quantify the dynamic impact of a regional shock in corporate default events over multiple periods. To do so, I separately conduct an impulse response analysis across all four regions, where the shock solely relates to the dynamic impact of a one period increase in each region's corporate default events on default risk exposure over multiple periods.

Before setting up the impulse response analysis, I first select a suitable bin size for each region. In the U.S. and Emerging Markets region, I focus the impulse response analysis on bin size of 6. For the Europe and Asia Pacific region, I focus on bin size 4. This approach allows us to construct an approximate average of 1.5 months in each rescaled interval period, allowing for a convenient comparison of impulse response analysis across different regions.

For the impulse response analysis in each region, I introduce a shock in period t = 0 that is equivalent to the theoretical mean of the bin size in each region.[15] In Figure 7, all regions except for the Asia Pacific Region, show that the shock on corporate default events is persistent. From the perspective of the U.S. economy, an increase of 6 default events in t = 0 will lead to an average increase of close to 3 default events in the next period. This impact, in turn, triggers additional default events over several periods in the future. A similar result is also observed in Europe and Emerging Market region. For the case of Europe, an increase of 4 default events in t = 0 leads to an increase of almost 2 default events in the next period.

To reiterate, the impulse response analysis confirms that corporate default events are contagious in each region. The economic intuition for contagious default events based on the impulse response analysis is similar to the Hawkes model, as documented by prior studies on corporate default contagion. Specifically, the impulse response analysis allows us to quantify the impact of one period shock in corporate default events on the increase

[15] Accordingly, as we assume that the data in time rescaled interval follows a Poisson distribution, a theoretical mean shock is equivalent to a theoretical variance shock. For the case of the U.S. and Emerging Markets, I consider a theoretical mean shock of 6 default events. For the case of Europe and Asia Pacific, a theoretical mean shock of 4 default events.

in subsequent default events over multiple periods. However, the key difference in our research lies in assessing default contagion based on time rescaled data. Our result suggests evidence of default contagion, which cannot be accounted for by firm-specific and systematic variables. This finding differs from the results that are documented by Lando and Nielsen (2010) and Azizpour et al. (2018), which solely focus on the U.S. economy.

As an additional robustness check, I also conduct another impulse response analysis based on larger bin sizes for all four regions. This time, bin size 8 for the U.S. and Emerging Markets region is selected. Bin size 6 is selected for Europe and Asia Pacific region. This approach aims to construct an approximate of $2-3$ months in each unit of time rescaled interval. Figure 8 plots the dynamic response of default events based on a one period theoretical mean shock of corporate default events. The economic intuition we may gain from the result in Figure 8 is similar to the impulse response plot we documented earlier.

6.2 Cross-Region Corporate Default Contagion

So far, the empirical results of this paper mainly focused on the severity of corporate default contagion contained within a regional scale. I next assess whether regional corporate default risk exposure could be impacted by corporate default events from an external region. To assess the potential of cross-region default contagion, I first provide a cross-correlation plot between corporate default events across different regions on a global scale. Figure 11 shows the cross-correlation plot between U.S. default events with the rest of the world and default events in Emerging Markets with the rest of the world.

In Figure 11, the corporate default events in the main economy/region of interests are tabulated in time rescaled intervals. In my case, this relates to either U.S. or Emerging Markets. Default events from an external region are tabulated accordingly based on the corresponding time rescaled interval of the main region.[16] Figure 11 shows that default events in the U.S. economy, under time rescaled interval of bin size 5, display strong degree of cross correlation with default events of regions in other parts of the world. The Emerging Markets region also depicts a similar observation. Specifically, corporate default events in the Emerging Markets region (under time rescaled interval) display strong degree of correlation with default events in the U.S. economy and Asia Pacific region over multiple lags. To reiterate, corporate default events in time rescaled intervals are supposed to follow a Poisson distribution. In this case, correlation in default events indicates the excess degree of dependence among default events across different regions that explanatory variables cannot account for. This observation raises preliminary concerns of cross-region corporate default contagion. Repeat of cross-correlation analysis at larger bin sizes also depict similar observations. However, we only report results for bin size 5 as this approach

[16] If certain time rescaled interval falls within the same date, corporate events are divided evenly within these periods

provides sufficient data count for additional econometrics analysis, which will be discussed in the later subsection. These results indicate that corporate default risk in the U.S. economy and emerging markets are vulnerable to default events in the rest of the world. [17]

6.2.1 Multivariate Poisson Autoregressive Model

To quantify the impact of cross-region corporate default contagion, I use a linear multivariate count autoregression model based on Fokianos et al. (2020). An unrestricted variant of the linear model is written as:

$$\lambda_t = d + A\lambda_{t-1} + BY_{t-1} \tag{9}$$

Both Y_t and λ_t are a vector. In the context of this paper, $\{Y_t = Y_{i,t}$, i = 1, 2, 3, 4 , $t \in \mathbb{Z}\}$, relates to a vector of default events. Y_{1t} relates to the main variable of accumulated corporate default events in time rescaled intervals. $\{Y_{i+1,t}, i = 1, 2, 3\}$ relates to the corporate default events in the rest of the world that contains accumulated default events based on the corresponding time rescaled interval of Y_{1t}. $\{\lambda_{i,t}$, i = 1, 2, 3, 4 , $t \in \mathbb{Z}\}$ relates to the expected default risk in each region. To reiterate, if our proportional hazard model (1) is correctly specified, Y_{1t} should follow an independent Poisson distribution. Correspondingly, this also means that $\lambda_{1,t}$ should also not be impacted by corporate default events from the rest of the world. d is a k $\times$ 1 vector, where the parameters are to be estimated. A, and B are k $\times$ k matrices, where the parameters are estimated.

However, due to the relative short time series of about 200 observations, I add several restrictions to (9) to avoid parameter proliferation. Specifically, I restrict the multivariate count autoregression to a bivariate scale and matrix A to a diagonal matrix. These additional restrictions are necessary due to the shorter time series, a commonly encountered issue in multivariate econometrics models that evaluates for contagion across financial variables.[18] A restricted variant of the multivariate Poisson Autoregressive model can be written as:

$$\begin{bmatrix}\lambda_{1,t}\\ \lambda_{2,t}\end{bmatrix} = \begin{bmatrix}d_1\\ d_2\end{bmatrix} + \begin{bmatrix}a_{11} & 0\\ 0 & a_{22}\end{bmatrix}\begin{bmatrix}\lambda_{1,t-1}\\ \lambda_{2,t-1}\end{bmatrix} + \begin{bmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{bmatrix}\begin{bmatrix}Y_{1,t-1}\\ Y_{2,t-1}\end{bmatrix} \tag{10}$$

[17] I also conduct a similar analysis from the perspective of Europe and Asia Pacific, but find less evidence that default clustering in these regions is attributed to default events from other parts of the world. These findings could be due to the relatively lower count of corporate default events in these aforementioned regions, making identification of cross-region default contagion more challenging. An alternative approach to identify default contagion can be derived from market-based variables, such as stock price or CDS ((See Jorion and Zhang (2007), Jorion and Zhang (2009), etc)). I leave this alternative approach for future research.

[18] Some examples of a similar restrictions in econometrics models are Aït-Sahalia et al. (2014), Fokianos et al. (2020), to name a few. Some studies mentioned above may even employ time series variables containing more data observation than ours.

After controlling for the impact of corporate default events and expected default risk in the last period, (10) allows us to evaluate if default events in other external regions can impact default risk ($\lambda_{1,t}$) which is the main economy/region of our concern. The key parameter that allows us to evaluate this external impact is b_{12}, which is the main parameter of our interests.

Following the model (10), I separately estimate the parameters using time rescaled data in the U.S. and Emerging Markets. As discussed earlier, default events in other regions are compiled based on the corresponding time rescaled interval of the main region of interest.

Table 13 and Table 14 present the parameter estimates for the U.S. and Emerging Markets based on bin size 5 for both regions. As the number of observations in time rescaled interval for the U.S. and Emerging Markets is 207 and 239, respectively, each time rescaled interval corresponds to about 1.5 months on average. Simultaneously, the parameter estimates of the bivariate Poisson Autoregressive model supports the autocorrelation of default events in time rescaled interval of U.S. and Emerging Market, as well as cross-correlation of default events in these regions with the rest of the world. From the U.S. economy perspective, we observe b_{11} and b_{12} to be largely significant for all three models at the 5% level. These results point out that the U.S. economy is still vulnerable to corporate default events from the rest of the world (based on significant b_{12}), even after controlling for default contagion from firms within the economy. In this case, the omission of default contagion from the rest of the world will lead to a distorted analysis of the relevant sources of default clustering in the U.S. economy. Notably, a_{11} is no longer significant after estimating the parameters of a Poisson Autoregressive model in a multivariate setting. This observation could be due to the relatively short collection of time series variables, which makes estimating additional parameters relating to the persistence of default events more challenging. Nonetheless, my parameter estimates suggest that domestic and external contagion are relevant sources of default clustering, as both b_{11} and b_{12} are mostly significant at the 5% level.

From an economic intuition perspective, Table 13 suggests that an increase in corporate default events by 10 in the Europe, Asia Pacific, and Emerging Markets will lead to an approximate increase in default events in the U.S. economy by 1, 2, 2, respectively. The time period is approximately 2 months.

Based on the results in Table 14, a similar intuition can also be interpreted for Emerging Markets perspective. Specifically, the parameters b_{11} and b_{12} are significant at least at the 5% level for the models relating to default events from the U.S. and Asia Pacific region. These results suggest that external default events from the U.S. and Asia Pacific also contribute to the clustering of default events in the Emerging Markets region. This observation holds, even after controlling for domestic default contagion among firm within the region. Intuitively, our findings point out that default events in the U.S. and Asia Pacific trigger additional default events in the Emerging Markets by 2 and 1 in the next approximate 2 months, respectively.

In summary, my earlier results point out that both autocorrelation and cross-correlation of default events (from external regions) contribute to the dispersion of default events. This finding is primarily supported by Figure 6 and Figure 11, as well as empirical results in Table 13 and Table 14. In the context of global corporate default clustering, my findings indicate that both domestic default contagion and default contagion from external regions contribute to the clustering of default events in a specified region. Omitting any of these factors may lead to a distorted analysis of the relevant sources of default clustering.

Incorporating supply chains or data on other forms of firm interlinkages may allow for the better identification of the mechanism that drives cross-region default contagion. Unfortunately, extensive and detailed high-quality data on global firm networks are unavailable over a sufficiently long time series. For instance, Factset's global supply chain data is mainly available since 2010 for most regions. Moreover, firms in financial distress often lack detailed data on its network relationship with other firms and even balance sheet data. As such, usage of firm network data may require us to give up data covering several of the earlier financial crisis and a large proportion of firm-level data, especially data on financially distressed firms, due to missing data. This approach may result in underestimating the severity of default clustering, which derails from the primary goal of the paper.[19]

7 Conclusion

The global economy has suffered through several severe episodes of corporate default clustering. Yet, existing literature on corporate default clustering has exclusively focused on the severity of default clustering in the U.S. economy.

Based on the time rescaling method, I show significant evidence of excess clustering in global corporate default events, which firm fundamentals and external macroeconomic conditions cannot explain. Systemic risk — severe coordinated waves of corporate distress drive the excess clustering of default events over multiple periods up to two years rather than random intermittent clustering of default events.

I identify significant evidence of international corporate default contagion as a novel source of corporate default clustering. Based on corporate default data in time rescaled interval, I first use a Poisson Autoregressive model to show a severe degree of corporate

[19]Another related field of research (e.g. Hertzel et al. (2008)) studies the impact of the supply chain on financial contagion based on stock prices or other equities-based information. This approach is unfeasible for my paper as their paper focuses on financial contagion based on market-based information instead of actual corporate default event, which is a different goal from my paper.

default contagion at a regional scale. An impulse response analysis allows us to quantify the dynamic impact of corporate default contagion over multiple periods. An extension of the Poisson Autoregressive model in a bivariate setting shows a considerable degree of cross-region corporate default contagion. Specifically, the empirical results show that the U.S. economy and Emerging Markets are vulnerable to default events in other parts of the world, indicating that regional corporate default risk is vulnerable to both default contagion within the region and default events from other parts of the world.

Finally, my research findings present pressing policy implications. As standard corporate vulnerability, corporate ratings, and distress risk models are primarily estimated based on firm-specific and systematic variables, excess clustering of default events suggests that existing distress risk assessment will have potentially underestimated vulnerabilities in global corporate debt. This concern is especially pertinent during crisis periods, which facilitates the clustering of corporate default events. My findings advocate the urgency in exploring alternative econometrics methods that account for correlated distress in global corporate debt, which can be a potential avenue for future research.

References

Daron Acemoglu, Vasco M Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. The network origins of aggregate fluctuations. *Econometrica*, 80(5):1977–2016, 2012.

Daron Acemoglu, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. Systemic risk and stability in financial networks. *American Economic Review*, 105(2):564–608, 2015.

Şenay Ağca and Oya Celasun. Sovereign debt and corporate borrowing costs in emerging markets. *Journal of International Economics*, 88(1):198–208, 2012.

Arianna Agosto, Giuseppe Cavaliere, Dennis Kristensen, and Anders Rahbek. Modeling corporate defaults: Poisson autoregressions with exogenous covariates (parx). *Journal of Empirical Finance*, 38:640–663, 2016.

Yacine Aït-Sahalia, Roger JA Laeven, and Loriana Pelizzon. Mutual excitation in eurozone sovereign cds. *Journal of Econometrics*, 183(2):151–167, 2014.

Edward I Altman. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, 23(4):589–609, 1968.

Pol Antràs and Davin Chor. Organizing the global value chain. *Econometrica*, 81(6):2127–2204, 2013.

Pol Antràs and Alonso De Gortari. On the geography of global value chains. *Econometrica*, 88 (4):1553–1598, 2020.

Pol Antras, Teresa C Fort, and Felix Tintelnot. The margins of global sourcing: Theory and evidence from us firms. *American Economic Review*, 107(9):2514–64, 2017.

Gonzalo Asis, Anusha Chari, and Adam Haas. In search of distress risk in emerging markets. *Journal of International Economics*, 131:103463, 2021.

Shahriar Azizpour, Kay Giesecke, and Gustavo Schwenkler. Exploring the sources of default clustering. *Journal of Financial Economics*, 129(1):154–183, 2018.

Mascia Bedendo and Paolo Colla. Sovereign and corporate credit risk: Evidence from the eurozone. *Journal of corporate Finance*, 33:34–52, 2015.

Ben S Bernanke, Mark Gertler, and Simon Gilchrist. The financial accelerator in a quantitative business cycle framework. *Handbook of macroeconomics*, 1:1341–1393, 1999.

Julia Bevilaqua, Galina B Hale, and Eric Tallman. Corporate yields and sovereign yields. *Journal of International Economics*, 124:103304, 2020.

Harjoat S Bhamra, Lars-Alexander Kuehn, and Ilya A Strebulaev. The levered equity risk premium and credit spreads: A unified framework. *The Review of Financial Studies*, 23(2): 645–703, 2010.

Francisco Blasques, Siem Jan Koopman, Andre Lucas, and Julia Schaumburg. Spillover dynamics for systemic risk measurement using spatial financial time series models. *Journal of Econometrics*, 195(2):211–223, 2016.

Audra L Boone and Vladimir I Ivanov. Bankruptcy spillover effects on strategic alliance partners. *Journal of Financial Economics*, 103(3):551–569, 2012.

John Y Campbell, Jens Hilscher, and Jan Szilagyi. In search of distress risk. *The Journal of Finance*, 63(6):2899–2939, 2008.

Bertrand Candelon and Sessi Tokpavi. A nonparametric test for granger causality in distribution with application to financial contagion. *Journal of Business & Economic Statistics*, 34(2):240–253, 2016.

Bertrand Candelon, Gilbert Colletaz, Christophe Hurlin, and Sessi Tokpavi. Backtesting value-at-risk: a gmm duration-based test. *Journal of Financial Econometrics*, 9(2):314–343, 2011.

Sudheer Chava and Robert A Jarrow. Bankruptcy prediction with industry effects. *Review of finance*, 8(4):537–569, 2004.

Sudheer Chava and Amiyatosh Purnanandam. Is default risk negatively related to stock returns? *The Review of Financial Studies*, 23(6):2523–2559, 2010.

Hui Chen, Winston Dou, Hongye Guo, and Yan Ji. Feedback and contagion through distressed competition. *The Rodney L. White Center Working Papers Series at the Wharton School, Jacobs Levy Equity Management Center for Quantitative Financial Research Paper*, 2020.

Anna Chernobai, Philippe Jorion, and Fan Yu. The determinants of operational risk in us financial institutions. *Journal of Financial and Quantitative Analysis*, 46(6):1683–1725, 2011.

Peter Christoffersen and Denis Pelletier. Backtesting value-at-risk: A duration-based approach. *Journal of Financial Econometrics*, 2(1):84–108, 2004.

John M Clapp, Yongheng Deng, and Xudong An. Unobserved heterogeneity in models of competing mortgage termination risks. *Real Estate Economics*, 34(2):243–273, 2006.

Antonio Coppola, Matteo Maggiori, Brent Neiman, and Jesse Schreger. Redrawing the map of global capital flows: The role of cross-border financing and tax havens. *The Quarterly Journal of Economics*, 136(3):1499–1556, 2021.

Sanjiv R Das, Darrell Duffie, Nikunj Kapadia, and Leandro Saita. Common failings: How corporate defaults are correlated. *The Journal of Finance*, 62(1):93–117, 2007.

Yongheng Deng, John M Quigley, and Robert Van Order. Mortgage terminations, heterogeneity and the exercise of mortgage options. *Econometrica*, 68(2):275–307, 2000.

Winston Dou, Shane A Johnson, MingMing Ao Shao, and Wei Wu. Competition network, distress propagation, and stock returns. *Distress Propagation, and Stock Returns (September 26, 2021)*, 2021.

Jin-Chuan Duan, Jie Sun, and Tao Wang. Multiperiod corporate default prediction—a forward intensity approach. *Journal of Econometrics*, 170(1):191–209, 2012.

Darrell Duffie, Andreas Eckner, Guillaume Horel, and Leandro Saita. Frailty correlated default. *The Journal of Finance*, 64(5):2089–2123, 2009.

Matthew Elliott, Benjamin Golub, and Matthew O Jackson. Financial networks and contagion. *American Economic Review*, 104(10):3115–53, 2014.

Konstantinos Fokianos, Anders Rahbek, and Dag Tjøstheim. Poisson autoregression. *Journal of the American Statistical Association*, 104(488):1430–1439, 2009.

Konstantinos Fokianos, Bård Støve, Dag Tjøstheim, and Paul Doukhan. Multivariate count autoregression. *Bernoulli*, 26(1):471–499, 2020.

Kay Giesecke. Correlated default with incomplete information. *Journal of Banking & Finance*, 28(7):1521–1545, 2004.

Kay Giesecke and Baeho Kim. Systemic risk: What defaults are telling us. *Management Science*, 57(8):1387–1405, 2011.

Heejoon Han, Oliver Linton, Tatsushi Oka, and Yoon-Jae Whang. The cross-quantilogram: Measuring quantile dependence and testing directional predictability between time series. *Journal of Econometrics*, 193(1):251–270, 2016.

Tarek Alexander Hassan, Jesse Schreger, Markus Schwedeler, and Ahmed Tahoun. Sources and transmission of country risk. Technical report, National Bureau of Economic Research, 2021.

Jean Helwege and Gaiyan Zhang. Financial firm bankruptcy and contagion. *Review of Finance*, 20(4):1321–1362, 2016.

Michael G Hertzel, Zhi Li, Micah S Officer, and Kimberly J Rodgers. Inter-firm linkages and the wealth effects of financial distress along the supply chain. *Journal of Financial Economics*, 87(2):374–387, 2008.

Philippe Jorion and Gaiyan Zhang. Good and bad credit contagion: Evidence from credit default swaps. *Journal of Financial Economics*, 84(3):860–883, 2007.

Philippe Jorion and Gaiyan Zhang. Credit contagion from counterparty risk. *The Journal of Finance*, 64(5):2053–2087, 2009.

Dimitris Karlis and Evdokia Xekalaki. A simulation comparison of several procedures for testing the poisson assumption. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 49(3):355–382, 2000.

Siem Jan Koopman, André Lucas, and Bernd Schwaab. Modeling frailty-correlated defaults using many macroeconomic covariates. *Journal of Econometrics*, 162(2):312–325, 2011.

Jun Hee Kwak. Corporate-sovereign debt nexus and externalities. *Available at SSRN 3741764*, 2020.

David Lando and Mads Stenbo Nielsen. Correlation in corporate defaults: Contagion or conditional independence? *Journal of Financial Intermediation*, 19(3):355–372, 2010.

Larry HP Lang and René M Stulz. Contagion and competitive intra-industry effects of bankruptcy announcements: An empirical analysis. *Journal of financial economics*, 32(1): 45–60, 1992.

Yanru Lee. Global corporate default risk factors: Frailty and spillover effects. *Working Paper*, 2022.

Oliver Linton and Yoon-Jae Whang. The quantilogram: With an application to evaluating directional predictability. *Journal of Econometrics*, 141(1):250–282, 2007.

Paul-André Meyer. Démonstration simplifiée d'un théorème de knight. *Séminaire de probabilités de Strasbourg*, 5:191–195, 1971.

NUS-CRI. NUS credit research initiative technical report. Technical report, National University of Singapore, 2021.

Frederick Dongchuhl Oh. Contagion of a liquidity crisis between two firms. *Journal of Financial Economics*, 107(2):386–400, 2013.

James A Ohlson. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of accounting research*, pages 109–131, 1980.

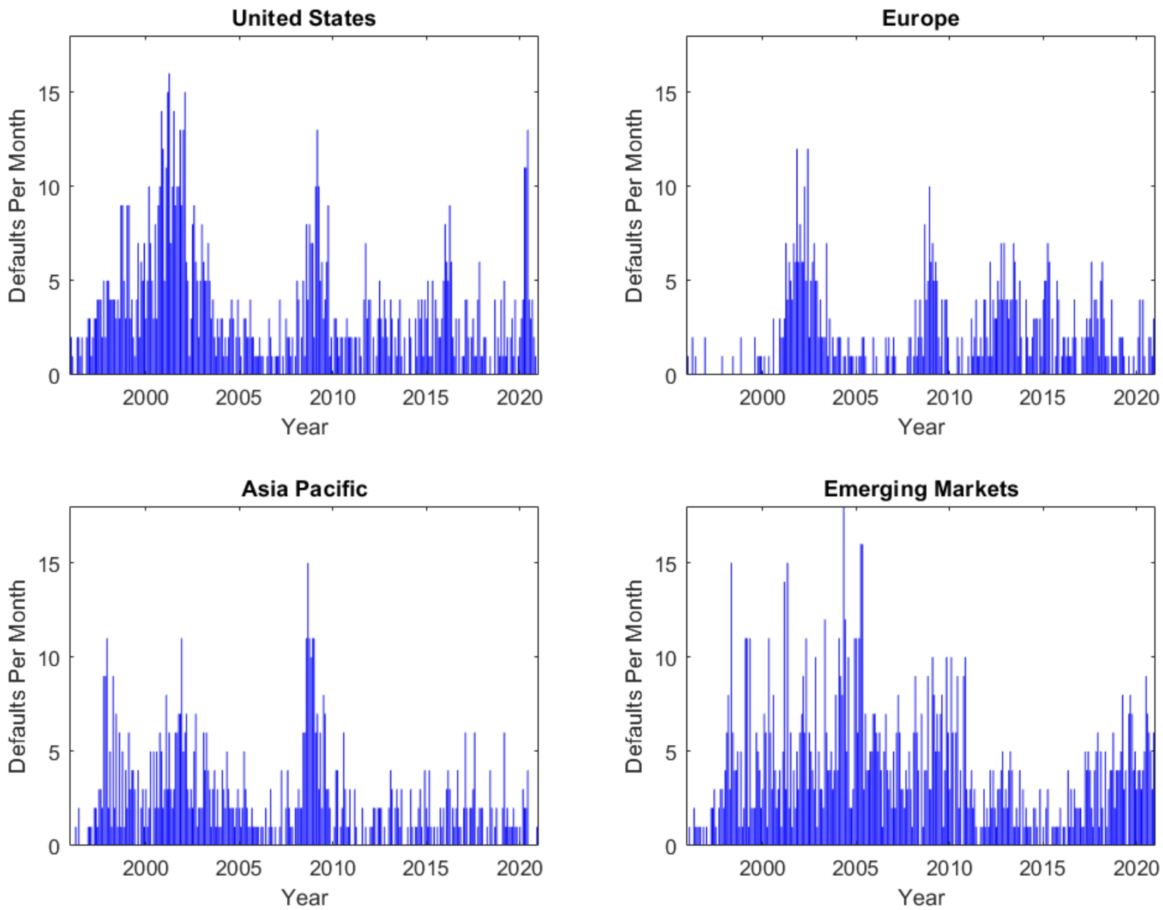
Gustavo Schwenkler and Hannan Zheng. The network of firms implied by the news. 2020.

Tyler Shumway. Forecasting bankruptcy more accurately: A simple hazard model. *The journal of business*, 74(1):101–124, 2001.

Steve Pak Yeung Wu. Corporate balance sheets and sovereign risk premia. *Available at SSRN 3685278*, 2020.

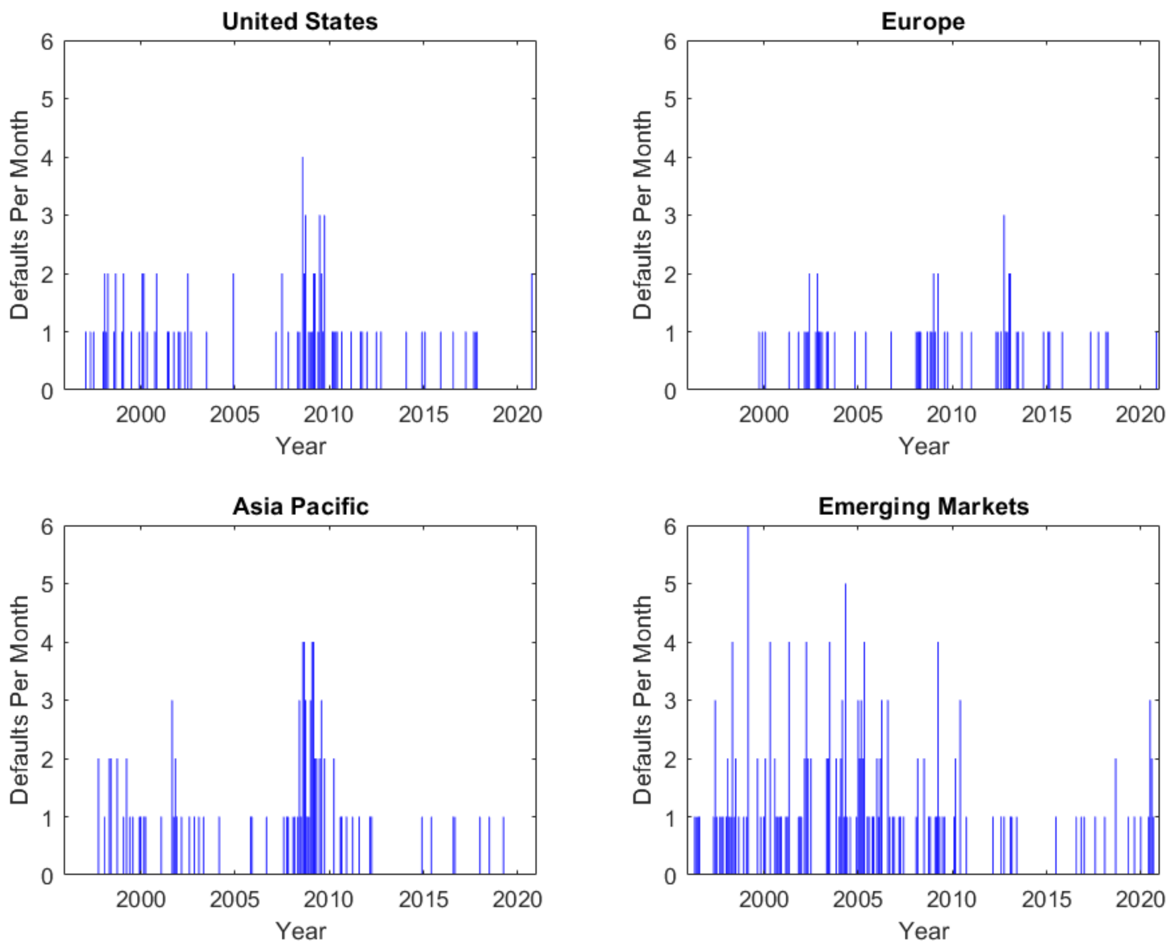
Mark E Zmijewski. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting research*, pages 59–82, 1984.

Figure 1: Global Corporate Default Events



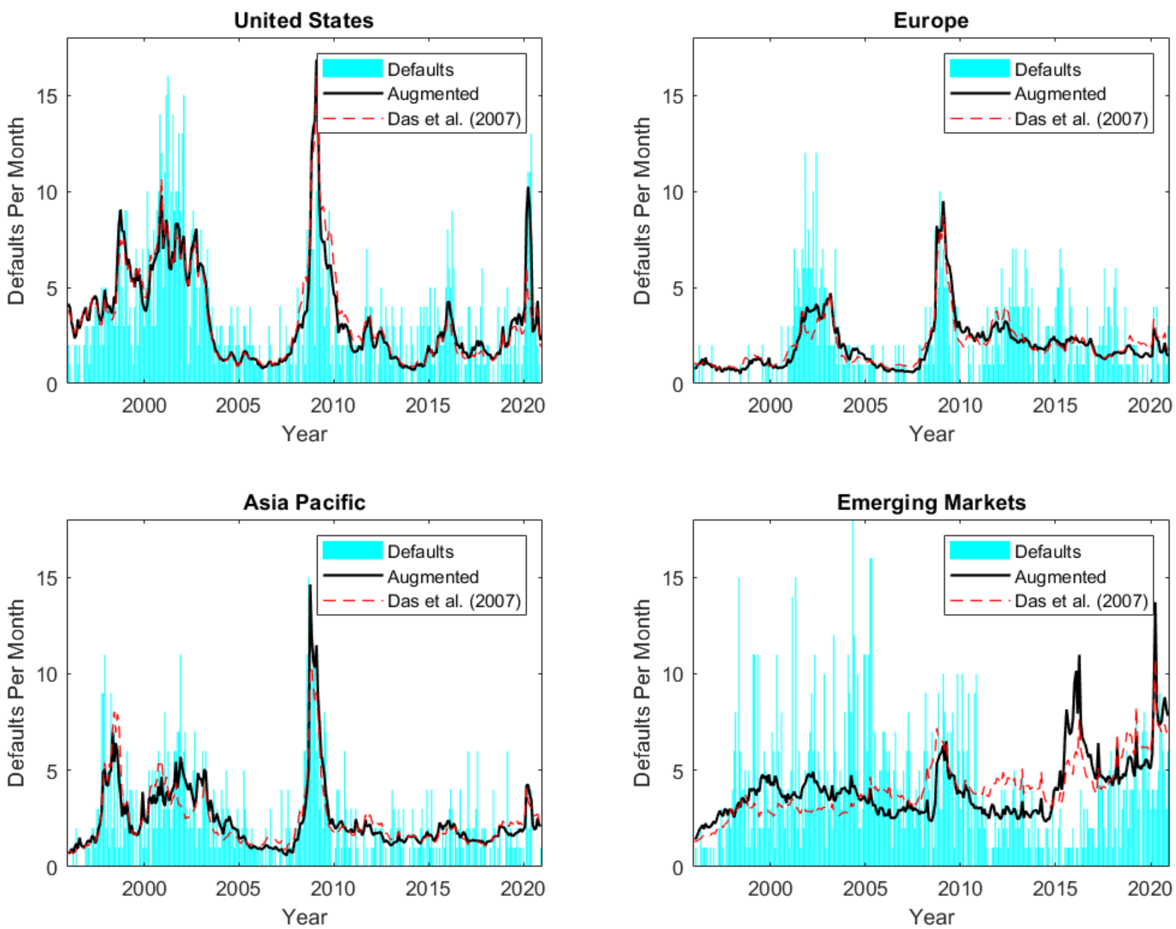
This figure shows monthly count of publicly listed corporate default events in four main countries/regions of the world: United States, Europe, Asia Pacific, and the Emerging Markets. Time period is from 1996 - 2020. Refer to table 1 and 2 for countries/economies in each region, and the definition of corporate default events. Data source: National University of Singapore Credit Research Initiative.

Figure 2: Global Corporate Default Events (Financial Firms only)



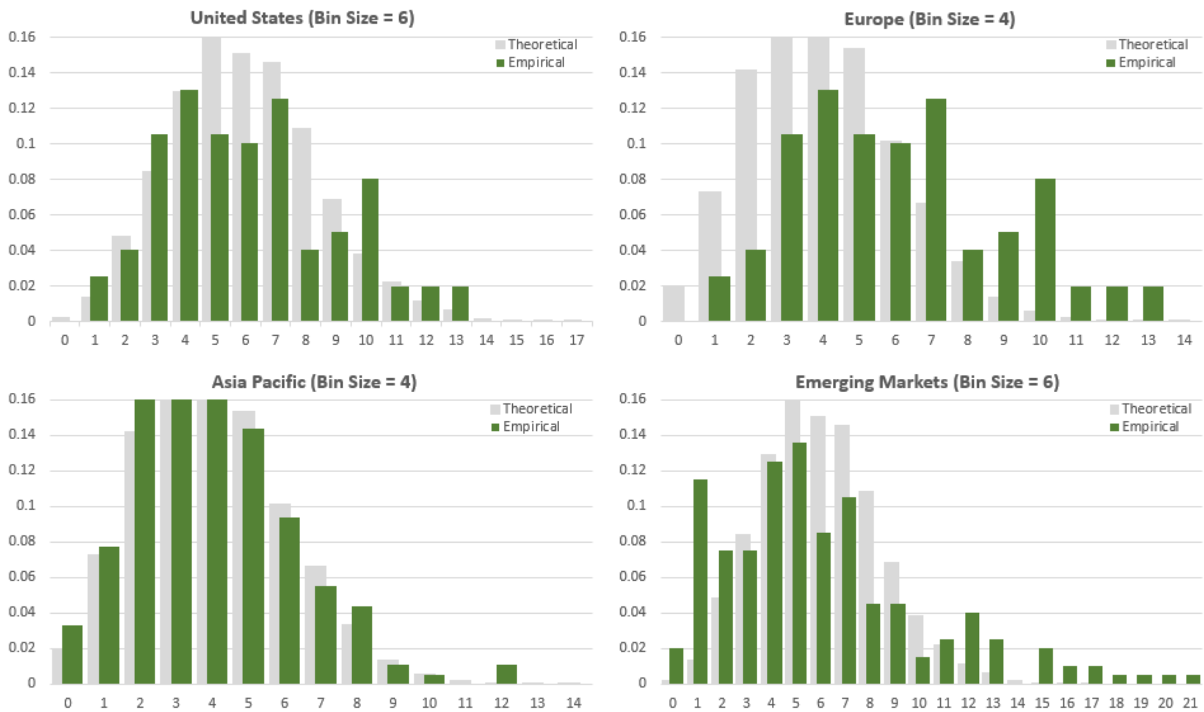
This figure shows monthly count of default events by financial firms in four main countries/regions of the world: United States, Europe, Asia Pacific, and the Emerging Markets. Time period is from 1996 - 2020. Refer to table 1 and 2 for countries/economies in each region, and the definition of corporate default events. Data source: National University of Singapore Credit Research Initiative.

Figure 3: Global Corporate Default Risk Prediction



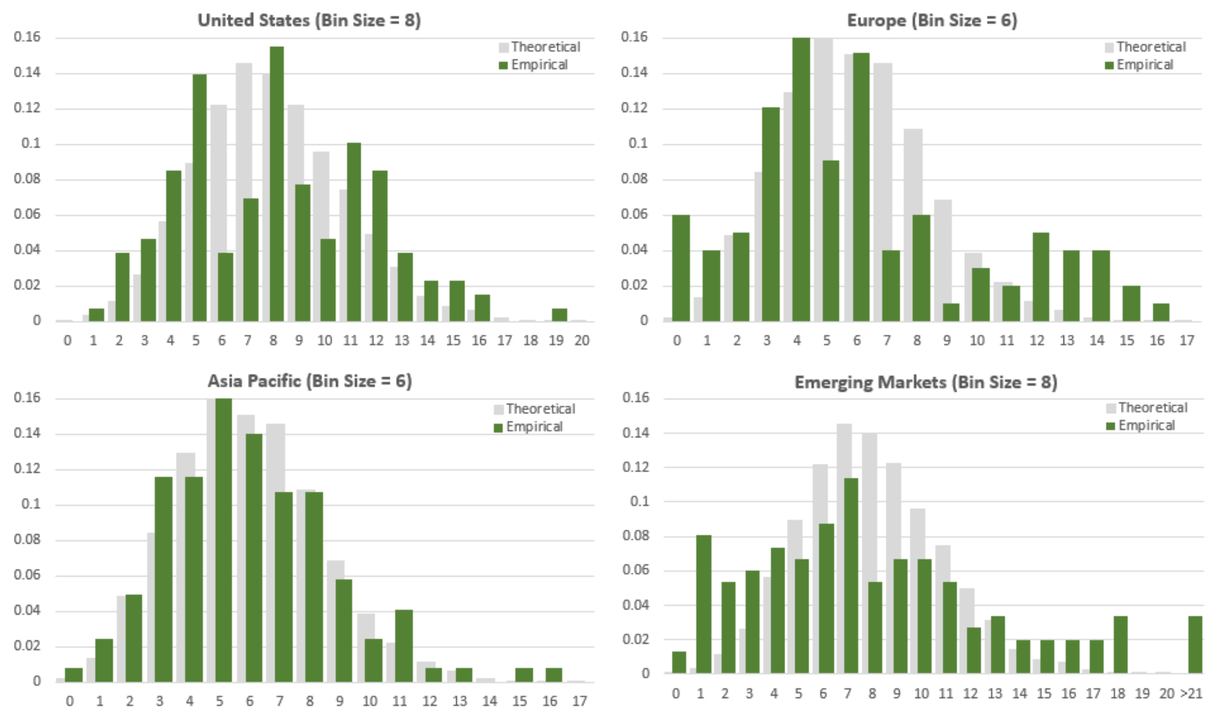
Time series comparison of actual and predicted default intensity (aggregate) for all four regions. This figure shows the actual number of defaults per month and the corresponding aggregate defaults predicted using the proportional hazard model as specified in (1). The number of predicted defaults in a month is the sum of the estimated probability of default for all firms. Dark solid line is aggregate default intensity for augmented model (all the explantory variables in Table 6), while dashed orange line is aggregate default intensity based on Das et al. (2007) explanatory variables.

Figure 4: Empirical vs Theoretical (Bin Sizes 4 and 6)



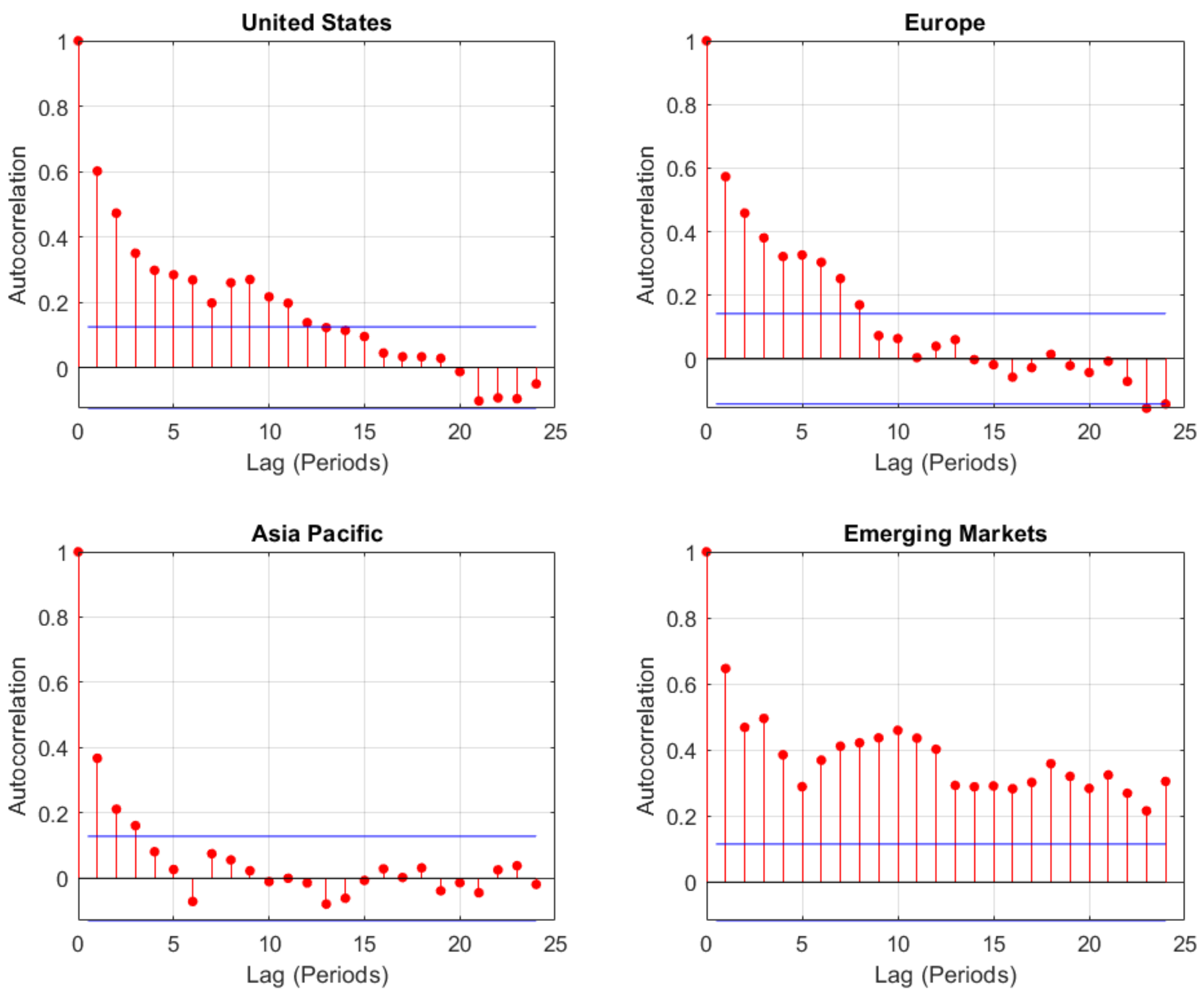
The Figure compares the distribution of global corporate default data in time rescaled interval and its theoretical Poisson distribution counterpart. Corporate default data in the U.S. and Emerging Markets are plotted based on Bin Size 6. Data for Europe and Asia Pacific are plotted based on Bin Size 4. Each region contains about 200 - 300 data observation. This corresponds to an average of 1 - 1.5 months in each time rescaled interval.

Figure 5: Empirical vs Theoretical (Bin Sizes 6 and 8)



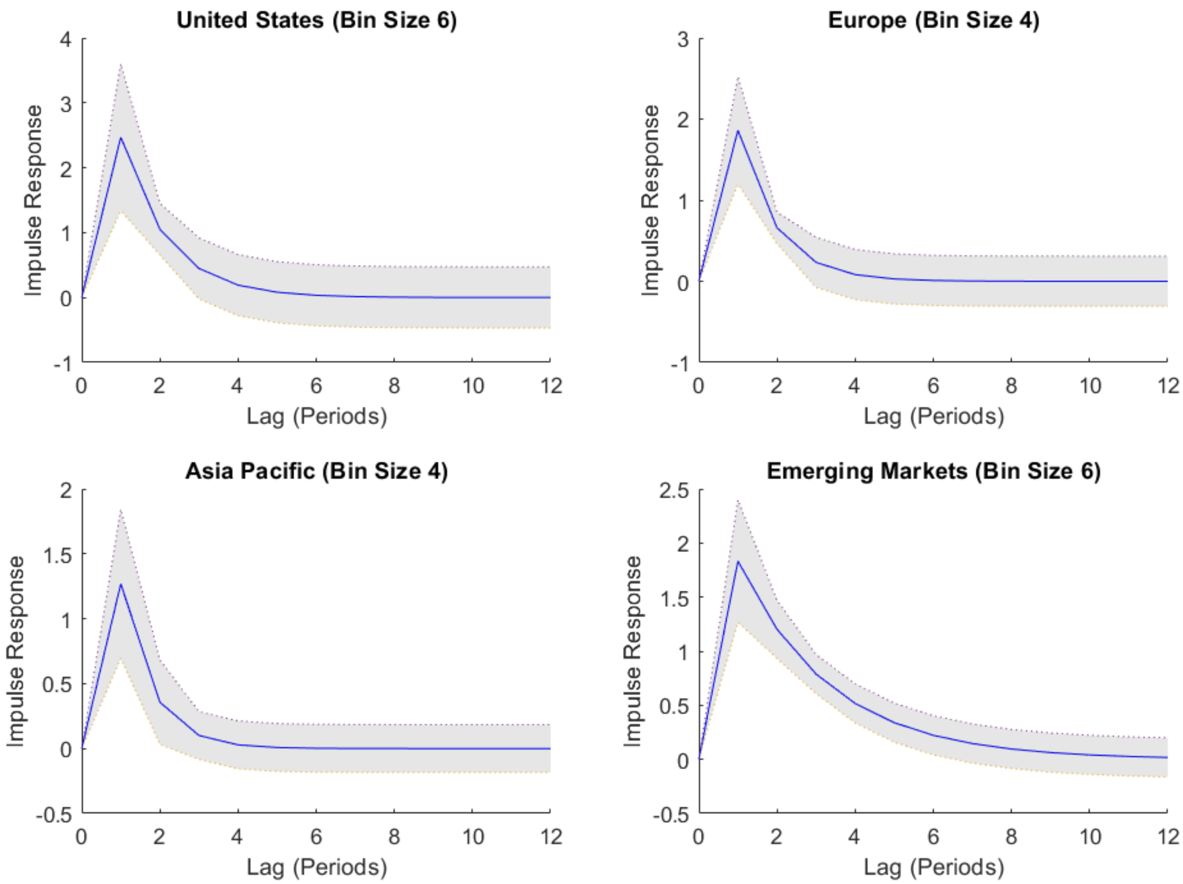
The Figure compares the distribution of global corporate default data in time rescaled interval and its theoretical Poisson distribution counterpart. Corporate default data in the U.S. and Emerging Markets are plotted based on Bin Size 6. Data for Europe and Asia Pacific are plotted based on Bin Size 4. Each region contains about 100 - 150 data observation. This corresponds to an average of 2 - 3 months in each time rescaled interval.

Figure 6: Global Corporate Default Autocorrelation in Time Rescaled Interval



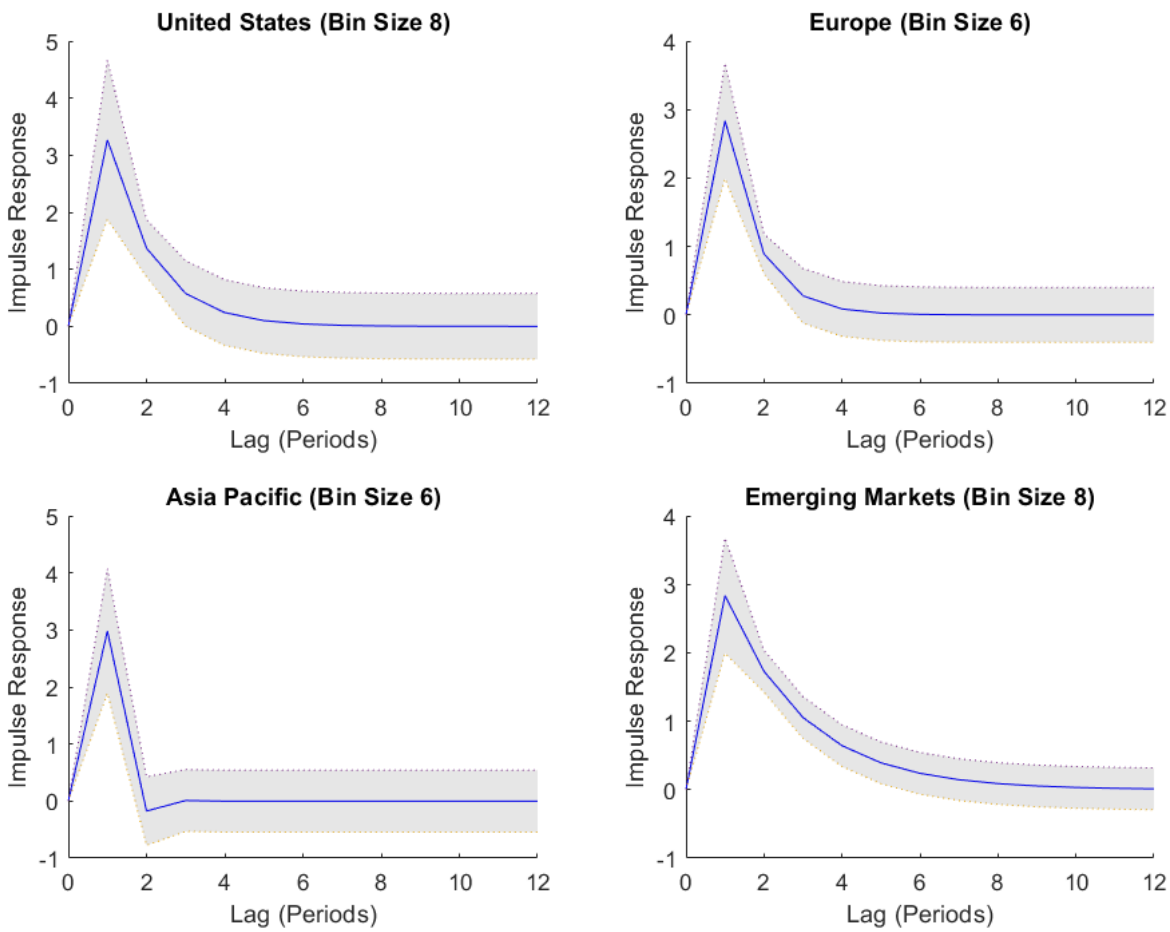
This figure provides a measure of autocorrelation among corporate default data in our time rescaled interval, across different lags. Data is presented for all regions. To prevent cluttering of information, we use bin size 5 for the United States and Emerging Markets. We use bin size 3 for Europe and Asia Pacific. Selection of other bin sizes also present similar results.

Figure 7: Regional Impulse Response Analysis



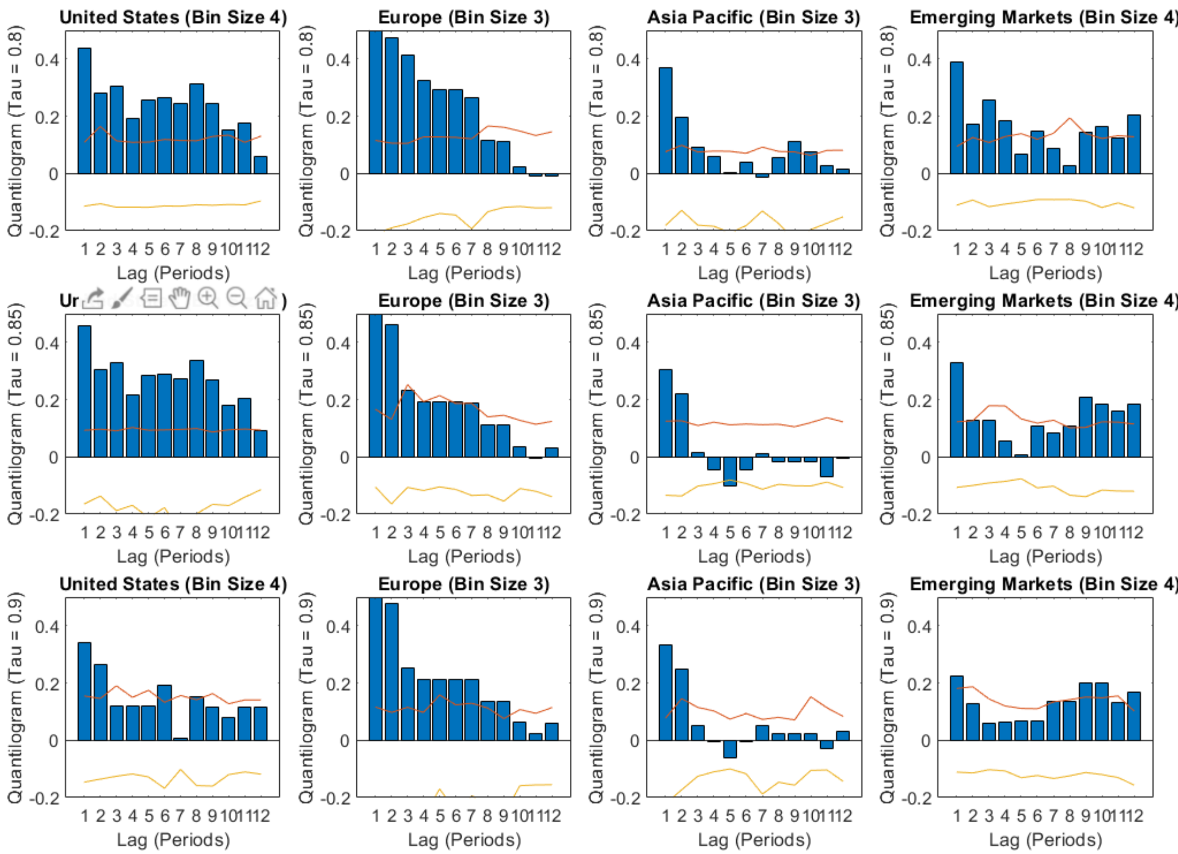
Regional impulse response analysis based on theoretical mean shock of bin size in each region. For the U.S. and Emerging Market, the theoretical mean shock is 6. For Europe and Asia Pacific, the theoretical mean shock is 4. Default data is in time rescaled interval. This correspond to 172, 149, 181, and 199 observations for the U.S., Europe, Asia Pacific, and Emerging Market region, respectively. One unit of data in each region is approximately an average of 1.5 months. Standard errors are computed based on delta method.

Figure 8: Regional Impulse Response Analysis (Robustness)



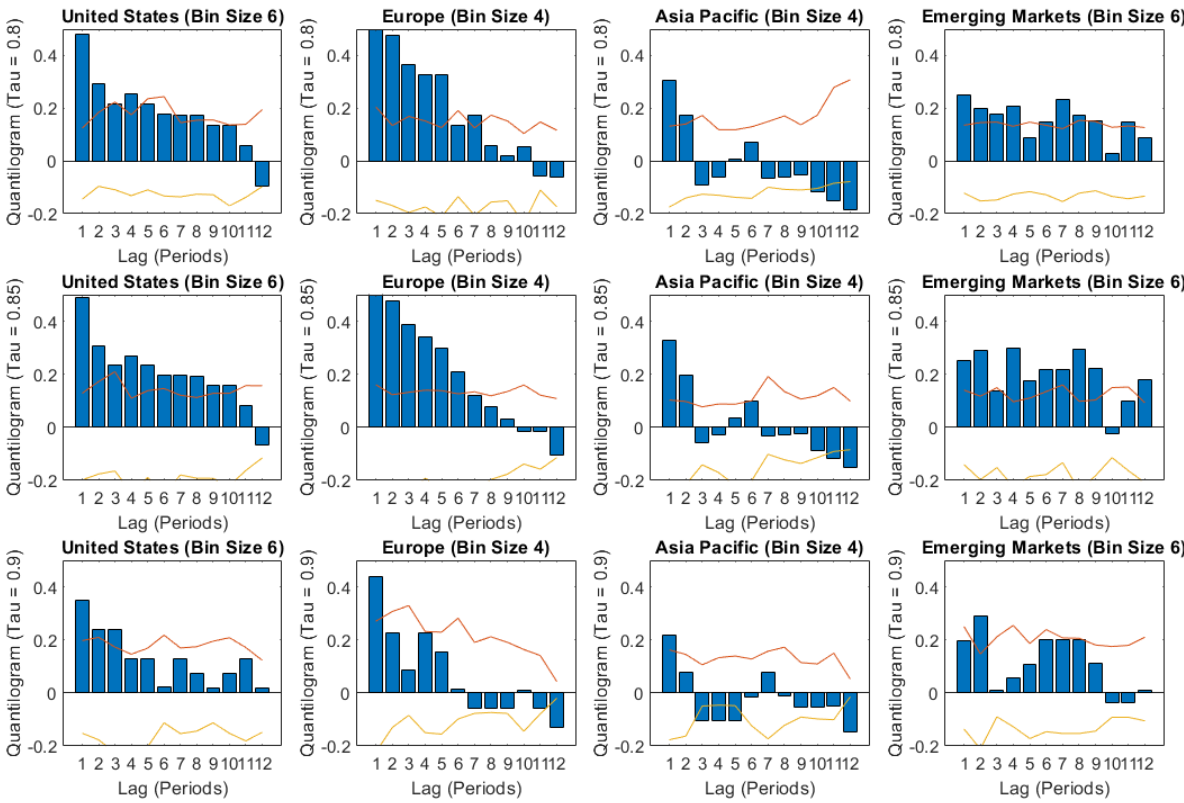
Regional impulse response analysis based on theoretical mean shock of bin size in each region. For the U.S. and Emerging Market, the theoretical mean shock is 8. For Europe and Asia Pacific, the theoretical mean shock is 6. Default data is in time rescaled interval. This correspond to 129, 99, 121, and 149 observations for the U.S., Europe, Asia Pacific, and Emerging Market region, respectively. One unit of data in each region is approximately an average of 2 - 3 months. Standard errors are computed based on delta method.

Figure 9: Quantilogram



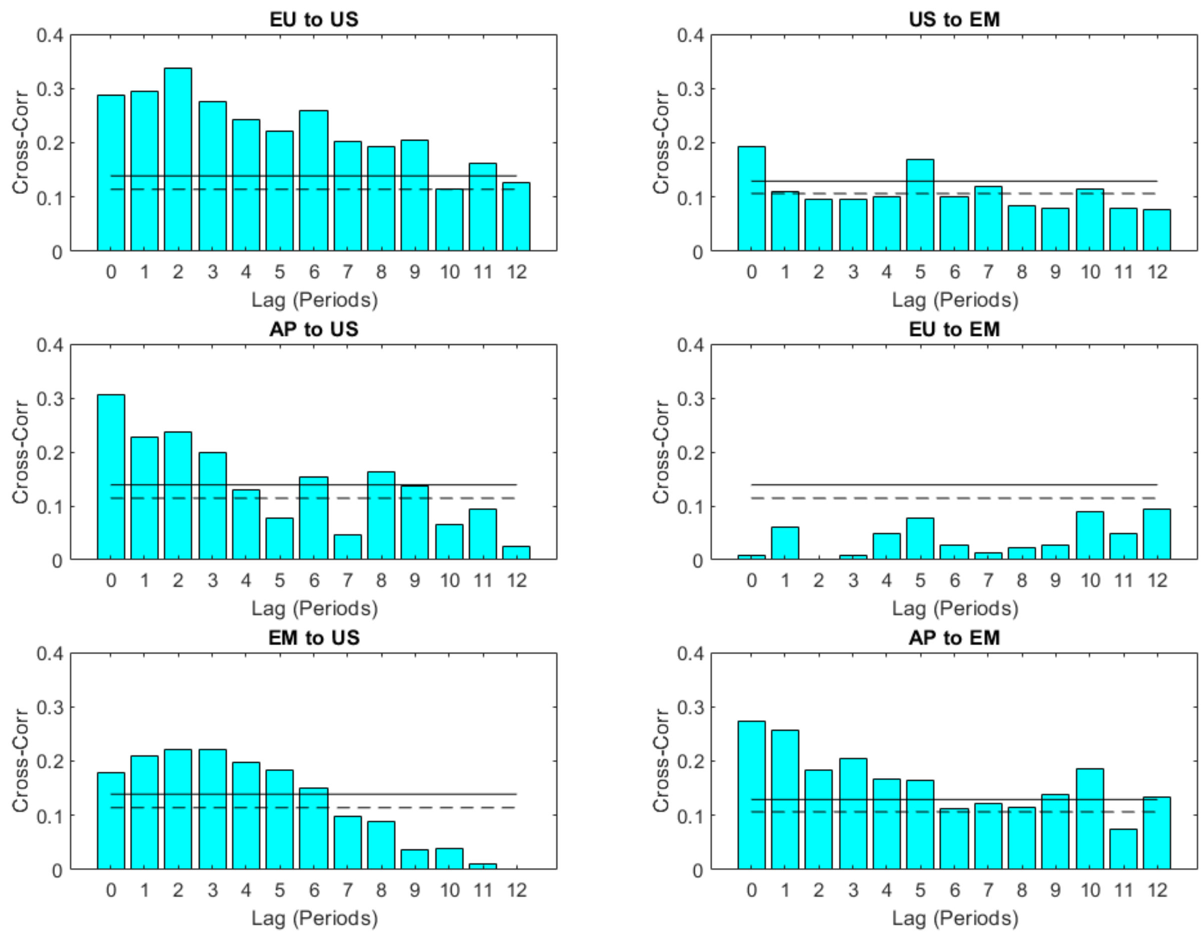
The figure shows the sample cross-quantilogram $\hat{p}(k)$ for the corporate default data of four regions at up to 12 lags. This correspond to 246, 195, 231, and 291 observations for the U.S., Europe, Asia Pacific, and Emerging Market region, respectively. One unit of data in each region is approximately an average of 1 - 1.5 months. Bar graphs describe the sample cross-quantilograms and lines are the 90% bootstrap confidence intervals centered at zero. Top panel correspond to τ =0.8, middle panel to τ =0.85, and low panel to τ =0.9.

Figure 10: Quantilogram



The figure shows the sample cross-quantilogram $\hat{p}(k)$ for the corporate default data of four regions at up to 12 lags. This correspond to 164, 147, 174, and 195 observations for the U.S., Europe, Asia Pacific, and Emerging Market region, respectively. One unit of data in each region is approximately an average of 1.5 - 2 months. Bar graphs describe the sample cross-quantilograms and lines are the 90% bootstrap confidence intervals centered at zero. Top panel correspond to τ =0.8, middle panel to τ =0.85, and low panel to τ =0.9.

Figure 11: Global Cross Correlation



The figure shows cross-correlation of default events. On the left column, default events in the U.S. economy are organized in time rescaled interval, based on section 3. Default events for the rest of the world are correspondinly compiled based on the time rescaled interval for the U.S. economy. The right column follows a similar approach, but from the perspective of the Emerging Markets. Bin Sizes are 5 for both U.S. and Emerging Markets (207 and 239 observations for U.S. and Emerging Markets respectively). Standard Error is plotted based on 95% confidence interval (Solid line), and 90% confidence interval (Dashed line).

Table 1: Regions: Classification of Countries/Economies

Regions	Economies/Countries
Developed Economies	
United States	-
Europe	United Kingdom, Norway, Poland, Switzerland, Austria, Belgium, Denmark, Finland, France, Germany, Greece, Italy, Netherlands, Portugal, Spain, Sweden
Asia Pacific	Australia, New Zealand,Japan, Hong Kong, Singapore, South Korea, Taiwan
Emerging Economies	
Emerging Markets	Mainland China, Indonesia, Malaysia, Philippines, Thailand, Argentina, Brazil, Colombia, Chile, Mexico

This table presents the classification of countries/economies into different regions/economic entities based on geographical proximity and similarities in structural characteristics of the economies

Table 2: Types of Corporate Default Events

Action Type	Subcategory
Bankruptcy	Administration, Arrangement, Canadian Companies' Creditors Arrangement Act (CCAA), Chapter 7,11,15 (United States bankruptcy code), Conservatorship, Insolvency, Japanese Corporate Reogranization Law (CRL), Judicial management, Liquidation, Pre-negotiation Chapter 11, Protection, Receivership, Rehabilitation, Rehabilitation (Thailand 1997), Reorganization, Restructuring, Section 304, Supreme Court declaration, Winding up, Workout, Sued by creditor, Petition withdrawn
Delisting	Bankruptcy
Default Corporate Action	Bankruptcy, Coupon & principal payment, Coupon payment only, Debt restructuring, Interest payment, Loan payment, Principal payment, Alternative Dispute Resolution (ADR, Japan only), Declared sick (India only), Regulatory action (Taiwan only), Financial difficulty and shutdown (Taiwan only), Buyback option

This table presents the three main types of corporate default events that are covered in the CRI database. Within each type of corporate default events, it can be further classified into numerous subcategories.

Table 3: Number and Default Rate per Economic Entities and Year

Year	United States			Europe			Asia Pacific			Emerging Markets		
	#Def	#Full	%	#Def	#Full	%	#Def	#Full	%	#Def	#Full	%
1996	15	72229	0.02	6	40591	0.01	4	55314	0.01	8	20481	0.04
1997	38	77001	0.05	1	45555	0	44	59189	0.07	19	24000	0.08
1998	57	77503	0.07	3	49840	0.01	44	62649	0.07	62	26147	0.24
1999	59	73604	0.08	5	52557	0.01	32	67156	0.05	59	27935	0.21
2000	89	71649	0.12	8	55839	0.01	38	71998	0.05	58	29020	0.2
2001	129	67566	0.19	64	60521	0.11	65	78528	0.08	61	29414	0.21
2002	78	62096	0.13	77	61075	0.13	45	85106	0.05	71	31274	0.23
2003	51	57687	0.09	27	58455	0.05	38	89867	0.04	59	33861	0.17
2004	28	55778	0.05	13	56998	0.02	30	94036	0.03	97	36421	0.27
2005	23	55825	0.04	10	57175	0.02	22	98775	0.02	95	39113	0.24
2006	13	55837	0.02	7	60103	0.01	11	104359	0.01	54	39730	0.14
2007	17	55169	0.03	8	65715	0.01	15	110295	0.01	50	41953	0.12
2008	52	54613	0.1	42	69556	0.06	81	115017	0.07	54	43713	0.12
2009	71	51728	0.14	46	67982	0.07	58	113547	0.05	74	44725	0.17
2010	22	49782	0.04	5	66874	0.01	26	114461	0.02	62	48492	0.13
2011	27	48947	0.06	24	66296	0.04	7	115668	0.01	21	53294	0.04
2012	29	47828	0.06	47	65418	0.07	13	116413	0.01	33	57311	0.06
2013	21	47029	0.04	44	63951	0.07	20	116849	0.02	30	59633	0.05
2014	25	47813	0.05	23	62824	0.04	17	118250	0.01	14	59508	0.02
2015	34	49167	0.07	40	62262	0.06	15	120752	0.01	13	59644	0.02
2016	44	48996	0.09	18	63719	0.03	21	124001	0.02	20	63753	0.03
2017	29	48567	0.06	32	64860	0.05	28	127575	0.02	40	69746	0.06
2018	13	49172	0.03	25	66331	0.04	12	130900	0.01	37	76764	0.05
2019	24	50133	0.05	9	66366	0.01	18	133795	0.01	59	80095	0.07
2020	58	50509	0.11	19	66017	0.03	13	136340	0.01	59	83637	0.07

This table presents the count of total default events, total number of firms months, and its respective default rate in each economic entities per year. Count of total firm months and default events are organized based on complete information as per our benchmark specification. Default events are counted based on next month default.

Table 4: Summary Statistics: Firm Specific Variables

U.S.	FullSample	Default	T-Test
DTD	3.811	-0.128	***
Stock Return Trail	1.372	-4.957	***
Profitability	-0.004	-0.090	***
Cash	0.081	0.072	**
Market to Book Ratio	2.227	2.407	**

Europe	FullSample	Default	T-Test
DTD	3.913	0.179	***
Stock Return Trail	1.024	-3.970	***
Profitability	-0.009	-0.048	***
Cash	0.085	0.082	
Market to Book Ratio	2.000	1.995	

Asia Pacific	FullSample	Default	T-Test
DTD	3.719	0.425	***
Stock Return Trail	1.404	-3.214	***
Profitability	-0.003	-0.030	***
Cash	0.136	0.071	***
Market to Book Ratio	1.437	1.500	†

Emerging Markets	FullSample	Default	T-Test
DTD	4.155	2.124	***
Stock Return Trail	0.774	-0.613	***
Profitability	-0.003	-0.008	***
Cash	0.081	0.043	***
Market to Book Ratio	1.999	2.525	***

Summary statistics for firm-months which contains data for all our selected firm-specific and systematic variables. The first two columns show simple means for full sample, and means for those firms that default next month. The last column shows the results of a two-sample t-test for equal means of each group of defaulted firms against the whole sample. ***, **, *, and † indicate $p < 0.01$, $p < 0.05$, $p < 0.10$, and $p < 0.15$

Table 5: Summary Statistics: Systematic Variables

U.S.	FullSample	Default	T-Test
Three Month Rate	2.428	2.328	*
Stock Index Trail	0.721	0.123	***
Yield Slope	0.762	0.808	*
Oil	50.456	44.199	***
Global Growth Rate	0.445	0.065	***

Europe	FullSample	Default	T-Test
Three Month Rate	2.132	1.863	***
Stock Index Trail	0.439	-0.434	***
Yield Slope	0.839	1.180	***
Oil	56.249	56.700	
Global Growth Rate	0.392	0.160	***

Asia Pacific	FullSample	Default	T-Test
Three Month Rate	1.650	3.543	***
Stock Index Trail	0.458	-0.648	***
Yield Slope	0.850	0.937	**
Oil	58.340	48.858	***
Global Growth Rate	0.376	0.119	***

Emerging Markets	FullSample	Default	T-Test
Three Month Rate	4.040	4.265	†
Stock Index Trail	0.498	0.229	***
Yield Slope	0.827	0.853	
Oil	58.979	50.794	***
Global Growth Rate	0.343	0.467	***

Summary statistics for firm-months which contains data for all our selected firm-specific and systematic variables. The first two columns show simple means for full sample, and means for those firms that default next month. The last column shows the results of a two-sample t-test for equal means of each group of defaulted firms against the whole sample. ***, **, *, and † indicate $p < 0.01$, $p < 0.05$, $p < 0.10$, and $p < 0.15$

Table 6: Parameter Estimates of Default Intensity Model (Global)

Parameters	United States	Europe	Asia Pacific	Emerging Markets
Constant	-4.808***	-6.331***	-6.196***	-5.554***
DTD	-2.434***	-1.223***	-1.599***	-0.426***
Stock Return (Trail)	-0.057***	-0.063***	-0.067***	-0.128
Three Month Rate	-0.123***	-0.057***	0.035***	-0.051***
Stock Index Return (Trail)	0.125***	-0.020	-0.030*	0.012
Profitability	-0.116***	0.012	0.006†	0.003
Liquidity	-0.383	-0.567*	-0.763**	-0.637†
Market to Book Ratio	0.128***	0.042**	0.197***	0.301***
Yield Slope	-0.070*	0.236***	0.232***	0.070**
Oil Price	-0.005***	-0.001	-0.004***	-0.009***
Global Growth Rate	-0.049***	-0.012	-0.020	-
Loglikelihood	-5730	-4301	-5208	-8894
AUC	0.966	0.914	0.926	0.791
Observations	1,426,228	1,516,880	2,560,840	1,179,664
Default	1,046	603	717	1,209

This table presents the parameter estimates of the proportional hazard for the U.S., Europe, Asia Pacific and Emerging Markets. The proportional hazard model includes both firm-specific variables and systematic variables. The dependent variable across all specification is binary, which indicates whether a firm has defaulted in the following month. We also included data on number of defaults events, firm observations, and likelihood. ***, **, *, and † indicate four levels of statistical significance of the coefficients: $p < 0.01$, $p < 0.05$, $p < 0.10$, and $p < 0.15$ respectively.

Table 7: Theoretical vs Empirical Moments (United States and Europe)

United States						
Type	BinSize	Count	Mean	Variance	Skewness	Kurtosis
Theory	2	508	2	2	0.71	3.5
Empirical	2	508	2.06	1.46	0.75	4.35
Theory	4	258	4	4	0.5	3.25
Empirical	4	258	4.08	4.19	0.47	2.71
Theory	6	172	6	6	0.41	3.17
Empirical	6	172	6.08	8.24	0.46	2.51
Theory	8	129	8	8	0.35	3.13
Empirical	8	129	8.09	13.29	0.3	2.55
Theory	10	104	10	10	0.32	3.1
Empirical	10	104	10.07	19.66	0.46	2.78
Theory	12	86	12	12	0.29	3.08
Empirical	12	86	12.12	27.02	0.39	2.89
Europe						
Type	BinSize	Count	Mean	Variance	Skewness	Kurtosis
Theory	2	296	2	2	0.71	3.5
Empirical	2	296	2.07	2.79	0.83	3.26
Theory	4	149	4	4	0.5	3.25
Empirical	4	149	4.03	8.04	0.91	3.33
Theory	6	99	6	6	0.41	3.17
Empirical	6	99	6.06	16.38	0.71	2.67
Theory	8	74	8	8	0.35	3.13
Empirical	8	74	8.04	26.53	0.98	3.22
Theory	10	59	10	10	0.32	3.1
Empirical	10	59	10.12	38.14	0.89	3.04
Theory	12	50	12	12	0.29	3.08
Empirical	12	50	12.08	53.1	0.81	2.73

This table presents the empirical and theoretical moments for default counts in each bin size: 2,4,6,8,10,12. For each of these bin size, we present results for the corresponding mean, variance, skewness and kurtosis. In part A of the table, we present the results for United States and Europe.

Table 8: Theoretical vs Empirical Moments (Asia Pacific and Emerging Markets)

Asia Pacific						
Type	BinSize	Count	Mean	Variance	Skewness	Kurtosis
Theory	2	359	2	2	0.71	3.5
Empirical	2	359	1.99	1.8	0.78	3.44
Theory	4	181	4	4	0.5	3.25
Empirical	4	181	3.94	5.06	0.75	3.8
Theory	6	121	6	6	0.41	3.17
Empirical	6	121	5.94	8.14	0.75	4
Theory	8	91	8	8	0.35	3.13
Empirical	8	91	7.86	12.86	0.79	3.88
Theory	10	72	10	10	0.32	3.1
Empirical	10	72	9.96	19.2	1.18	5.68
Theory	12	60	12	12	0.29	3.08
Empirical	12	60	11.92	20.96	0.55	3.36
Emerging Markets						
Type	BinSize	Count	Mean	Variance	Skewness	Kurtosis
Theory	2	589	2	2	0.71	3.5
Empirical	2	589	2.03	2.93	1.4	5.56
Theory	4	298	4	4	0.5	3.25
Empirical	4	298	4.07	9.33	1.37	5.48
Theory	6	199	6	6	0.41	3.17
Empirical	6	199	6.07	19.29	1.21	4.45
Theory	8	149	8	8	0.35	3.13
Empirical	8	149	8.07	31.62	1.31	5.82
Theory	10	120	10	10	0.32	3.1
Empirical	10	120	10.08	46.72	1.24	5.49
Theory	12	100	12	12	0.29	3.08
Empirical	12	100	12.1	61.34	0.92	3.89

This table presents the empirical and theoretical moments for default counts in each bin size: 2,4,6,8,10,12. For each of these bin size, we present results for the corresponding mean, variance, skewness and kurtosis. In part B of the table, we present the results for Asia Pacific and Emerging Markets.

Table 9: Global Econometrics Tests for Excess Default Clustering

United States						
Bins Size	Count	Fisher Dispersion	Upper Tail (Mean)	Upper Tail (Median)	PWB	KK
Bin Size 2	508	1.000	0.996	0.069*	0.000***	0.000***
Bin Size 4	258	0.284	0.248	0.000***	0.758	0.689
Bin Size 6	172	0.001***	0.001***	0.000***	0.001***	0.000***
Bin Size 8	129	0.000***	0.007***	0.001***	0.000***	0.004***
Bin Size 10	104	0.000***	0.000***	0.000***	0.000***	0.019**
Bin Size 12	86	0.000***	0.000***	0.000***	0.000***	0.129†
Europe						
Bins Size	Count	Fisher Dispersion	Upper Tail (Mean)	Upper Tail (Median)	PWB	KK
Bin Size 2	296	0.000***	0.000***	0.000***	0.000***	0.000***
Bin Size 4	149	0.000***	0.000***	0.000***	0.000***	0.000***
Bin Size 6	99	0.000***	0.000***	0.000***	0.000***	0.000***
Bin Size 8	74	0.000***	0.000***	0.000***	0.000***	0.000***
Bin Size 10	59	0.000***	0.000***	0.000***	0.000***	0.000***
Bin Size 12	50	0.000***	0.000***	0.000***	0.000***	0.000***
Asia Pacific						
Bins Size	Count	Fisher Dispersion	Upper Tail (Mean)	Upper Tail (Median)	PWB	KK
Bin Size 2	359	0.914	0.211	0.000***	0.190	0.000***
Bin Size 4	181	0.009***	0.201	0.000***	0.007***	0.000***
Bin Size 6	121	0.006***	0.039**	0.032**	0.004***	0.000***
Bin Size 8	91	0.000***	0.078*	0.163	0.000***	0.000***
Bin Size 10	72	0.000***	0.017**	0.020**	0.000***	0.000***
Bin Size 12	60	0.000***	0.116†	0.063*	0.000***	0.384
Emerging Markets						
Bins Size	Count	Fisher Dispersion	Upper Tail (Mean)	Upper Tail (Median)	PWB	KK
Bin Size 2	589	0.000***	0.000***	0.057*	0.000***	0.000***
Bin Size 4	298	0.000***	0.001***	0.000***	0.000***	0.000***
Bin Size 6	199	0.000***	0.000***	0.000***	0.000***	0.000***
Bin Size 8	149	0.000***	0.000***	0.000***	0.000***	0.000***
Bin Size 10	120	0.000***	0.000***	0.000***	0.000***	0.000***
Bin Size 12	100	0.000***	0.000***	0.000***	0.000***	0.000***

The table presents the p-values for series of statistical tests that are conducted on corporate default data allocated into time rescaled bins, which corresponds to the Poisson distribution. The statistical tests include: Fisher Dispersion (FD), Upper Tail Mean (UT1), Upper Tail Median (UT2), Potthoff-Whittinghill-Bohning (PWB), Kocherlakota-Kocherlakota (KK). ***, **, and * indicate three levels of statistical significance for the coefficients: $p < 0.01$, $p < 0.05$, and $p < 0.10$ respectively. Statistical results are presented across different regions: United States, Europe, Asia Pacific and Emerging Markets.

Table 10: Global Serial Correlation

United States					
Binsize	Count	A	T Stats (A)	B	T Stats (B)
Bin Size 2	508	0.85***	-12.67	0.59***	20.81
Bin Size 4	258	1.58***	-9.74	0.61***	12.35
Bin Size 6	172	2.5***	-7.57	0.59***	9.56
Bin Size 8	129	3.4***	-6.24	0.58***	7.4
Bin Size 10	104	4.28***	-5.87	0.57***	7.25
Bin Size 12	86	4.99***	-6.01	0.59***	7.35
Europe					
Binsize	Count	A	T Stats (A)	B	T Stats (B)
Bin Size 2	296	0.84***	-7.74	0.59***	12.64
Bin Size 4	149	1.58***	-6.25	0.61***	9.73
Bin Size 6	99	2.43***	-4.88	0.6***	7.02
Bin Size 8	74	3.23***	-3.65	0.6***	5.74
Bin Size 10	59	3.98***	-3.66	0.6***	5.39
Bin Size 12	50	5.21***	-3.04	0.57***	4.06
Asia Pacific					
Binsize	Count	A	T Stats (A)	B	T Stats (B)
Bin Size 2	359	1.05***	-7.54	0.47***	10.6
Bin Size 4	181	2.45***	-5.26	0.38***	6.86
Bin Size 6	121	3.08***	-5.86	0.48***	7.59
Bin Size 8	91	4.15***	-4.82	0.47***	6.01
Bin Size 10	72	7.34**	-2.04	0.26***	2.55
Bin Size 12	60	8.12***	-2.25	0.31***	2.48
Emerging Markets					
Binsize	Count	A	T Stats (A)	B	T Stats (B)
Bin Size 2	589	0.54***	-14.98	0.73***	30.25
Bin Size 4	298	1.42***	-9.19	0.65***	17.38
Bin Size 6	199	2.52***	-6.08	0.58***	10.9
Bin Size 8	149	3.26***	-6.34	0.59***	12
Bin Size 10	120	3.77***	-6.21	0.62***	10.9
Bin Size 12	100	4.25***	-5.42	0.65***	8.14

This table presents results for parameter estimates of autoregressive model in equation (3) across different bin sizes: 2,4,6,8,10,12. The t-statistics is shown on the right of the parameter coefficients. We test for serial correlation of default counts in each successive bins. That is, under the null hypothesis of doubly stochastic defaults, fixing an accumulative total default intensity of c per time bin, the numbers of defaults X1, X2, ... , XK in successive bins are independent and identically distributed. The parameter A is the intercept in the AR1 model and B is the autoregression coefficient. T-statistics for A are presented for the test A = c, based on respective bin size and not A =0. Statistical results are presented across different regions: United States, Europe, Asia Pacific and Emerging Markets. ***, **, *, and † indicate four levels of statistical significance of the coefficients: $p < 0.01$, $p < 0.05$, $p < 0.10$, and $p < 0.15$ respectively.

Table 11: Regional Corporate Default Contagion

United States							
Binsize	Count	ω	T Stats (ω)	α	T Stats (α)	β	T Stats (β)
Bin Size 2	508	0.36***	-13.01	0.49***	6.26	0.33***	2.81
Bin Size 4	258	0.74***	-10.77	0.46***	5.56	0.36***	2.77
Bin Size 6	172	1.01***	-10.74	0.41***	4.27	0.43***	2.86
Bin Size 8	129	1.44***	-10.94	0.41***	4.6	0.42***	3.14
Bin Size 10	104	1.88***	-11.51	0.42***	4.68	0.40***	3.23
Bin Size 12	86	2.65***	-10.62	0.46***	5.15	0.33***	2.76
Europe							
Binsize	Count	ω	T Stats (ω)	α	T Stats (α)	β	T Stats (β)
Bin Size 2	296	0.32***	-11.42	0.39***	4.75	0.46***	3.34
Bin Size 4	149	0.73***	-11.2	0.46***	5.5	0.35***	2.73
Bin Size 6	99	1.3***	-11.14	0.47***	6.59	0.31***	3.01
Bin Size 8	74	1.99***	-9.48	0.56***	6.73	0.19*	1.73
Bin Size 10	59	3.34***	-7.35	0.58***	6.67	0.09	0.76
Bin Size 12	50	4.75***	-5.73	0.55***	6.35	0.06	0.46
Asia Pacific							
Binsize	Count	ω	T Stats (ω)	α	T Stats (α)	β	T Stats (β)
Bin Size 2	359	1.46**	-1.77	0.51***	7.82	-0.25	-1.56
Bin Size 4	181	1.58***	-4.11	0.32***	4.33	0.28†	1.57
Bin Size 6	121	3.32***	-2.67	0.5***	5.35	-0.06	-0.32
Bin Size 8	91	7.64	-0.24	0.57***	5.99	-0.55***	-3.47
Bin Size 10	72	6.49*	-1.79	0.3***	2.76	0.05	0.21
Bin Size 12	60	7.55***	-2.18	0.33***	2.94	0.04	0.21
Emerging Markets							
Binsize	Count	ω	T Stats (ω)	α	T Stats (α)	β	T Stats (β)
Bin Size 2	589	0.3***	-13.44	0.64***	6.87	0.21†	1.46
Bin Size 4	298	0.22***	-32.7	0.38***	4.95	0.56***	5.86
Bin Size 6	199	0.25***	-51.68	0.31***	6.34	0.66***	11.98
Bin Size 8	149	0.3***	-49.85	0.35***	6.61	0.61***	10.38
Bin Size 10	120	0.48***	-42.92	0.38***	6.9	0.57***	9.25
Bin Size 12	100	0.75***	-37.18	0.43***	7.13	0.51***	7.38

This table presents results for the parameter estimates of the Poisson Autoregression (PAR) model in equation 8 across different regions and bin sizes. The t-statistics is shown on the right for each of the parameter coefficients. We test for corporate default contagion over multiple time period. The parameter ω is the intercept in the PAR model, α accounts for last period default events on period t default exposure, β measures the persistency of default events. T-statistics for ω are presented for the test ω = c, based on respective bin size and not $\omega = 0$. Statistical results are presented for the four regions: United States, Europe, Asia Pacific and Emerging Markets. ***, **, *, and † indicate four levels of statistical significance of the coefficients: $p < 0.01$, $p < 0.05$, $p < 0.10$, and $p < 0.15$ respectively.

Table 12: Tail Risk Serial Dependence

Bin Size	United States		Europe		Asia Pacific		Emerging Market	
	J Stat	P Value	J Stat	P Value	J Stat	P Value	J Stat	P Value
Bin Size 2	71.5	0.000	92.5	0.000	74.7	0.000	148.9	0.000
Bin Size 4	44.1	0.000	42.9	0.000	28.3	0.000	95.0	0.000
Bin Size 6	66.3	0.000	36.8	0.000	26.1	0.000	72.5	0.000
Bin Size 8	41.2	0.000	23.7	0.000	14.6	0.000	50.9	0.000
Bin Size 10	46.3	0.000	20.6	0.000	14.4	0.000	56.9	0.000
Bin Size 12	29.7	0.000	19.5	0.000	9.0	0.003	43.2	0.000

In this table, we present the J statistics, which converge to a Chi-square distribution with degree of freedom 1. We calculate the corresponding p-value that is computed based on the J-statistics. For this assignment, we consider the case of first order orthonormal polynomial. In Candelon et al. (2011), they show that the results largely remain the same, even after considering higher order orthonormal polynomial.

Table 13: Parameter Estimates of Bivariate Poisson Autoregressive Model
(Time Rescaled U.S. Corporate Default Data)

Parameters	EU ⇒ US	AP ⇒ US	EM ⇒ US
d_1	0.373***	0.332***	0.285***
	(0.313)	(0.298)	(0.242)
d_2	0.082***	0.232***	0.127***
	(0.071)	(0.121)	(0.117)
a_{11}	0.023	0.019	0.080
	(0.490)	(0.572)	(0.432)
a_{22}	0.539***	0.494***	0.519***
	(0.137)	(0.088)	(0.093)
b_{11}	0.818**	0.744**	0.686**
	(0.360)	(0.302)	(0.280)
b_{22}	0.377***	0.300**	0.387***
	(0.089)	(0.141)	(0.049)
b_{12}	0.079***	0.173*	0.185**
	(0.024)	(0.096)	(0.075)
b_{21}	-0.088	0.047	0.002
	(0.297)	(0.185)	(0.123)
Log Likelihood	348	348	350
AIC	-3.53	-3.03	-2.36
BIC	-3.37	-2.87	-2.20

This table presents the parameter estimates of the Bivariate Poisson Autoregressive Model. U.S. corporate default events are compiled based on time rescaled interval. Default events for the rest of world are correspondingly compiled based on time rescaled interval of U.S. default events. We also included data on log likelihood, AIC, and BIC. ***, **, *, and † indicate four levels of statistical significance of the coefficients: $p < 0.01$, $p < 0.05$, $p < 0.10$, and $p < 0.15$ respectively.

Table 14: Parameter Estimates of Bivariate Poisson Autoregressive Model
(Time Rescaled EM Corporate Default Data)

Parameters	US ⇒ EM	EU ⇒ EM	AP ⇒ EM
d_1	0.155***	0.246***	0.224***
	(0.144)	(0.086)	(0.169)
d_2	0.095***	0.104***	0.139***
	(0.060)	(0.106)	(0.056)
a_{11}	0.182	0.417***	0.089
	(0.365)	(0.091)	(0.275)
a_{22}	0.636***	0.106	0.559***
	(0.046)	(0.14)	(0.064)
b_{11}	0.795**	0.581***	0.784***
	(0.371)	(0.085)	(0.196)
b_{22}	0.230	0.836***	0.311***
	(0.166)	(0.159)	(0.097)
b_{12}	0.180**	0.103	0.117**
	(0.072)	(0.071)	(0.047)
b_{21}	-0.115	-0.045	0.058
	(0.305)	(0.128)	(0.158)
Log Likelihood	385	328	306
AIC	-3.95	-2.20	-1.79
BIC	-3.80	-2.06	-1.64

This table presents the parameter estimates of the Bivariate Poisson Autoregressive Model. Emerging Markets corporate default events are compiled based on time rescaled interval. Default Events for the rest of world are correspondingly compiled based on time rescaled interval of Emerging Markets default events. We also included data on log likelihood, AIC, and BIC. ***, **, *, and † indicate four levels of statistical significance of the coefficients: $p < 0.01$, $p < 0.05$, $p < 0.10$, and $p < 0.15$ respectively.

Appendix

Global Corporate Default Clustering and Contagion

Yanru Lee

A Appendix A

Table A.1: Variables Construction

Variable Name	Variables Construction
Trailing Stock Return	Trailing 1-year stock return
Profitability	Ratio of net income to the market value of total assets, where the market value of assets is equal to the sum of the firm's market capitalization and total liabilities
Cash	Ratio of cash and cash equivalents to the market value of total assets
Market to Book Ratio	Ratio of market capitalization to book value of equity, where book value of equity is total assets minus total liabilities. We follow Campbell et al. (2008) and Asis et al. (2021) in the calculation of this value. If a firm has negative book value of equity, the book value of equity is set to $1 so as to place that firm's market-to-book ratio in the right-hand side of the distribution
Distance to Default (DTD)	DTD provides a measure of firm's 'distance' to default. This value is computed based on firm's market value asset, volatility of asset, total liabilities, short and long term debt. Our DTD calculation is downloaded from NUS CRI and does not follow a standard calculation. Refer to NUS-CRI (2021) for more details.
Three Month Rate	Three month rate of the corresponding economy. This variable is from NUS CRI. Refer to NUS-CRI (2021) for more details.
Trailing Stock Indices	Trailing 1-year stock indices. Respective stock indices for an economy is based on NUS CRI. Refer to NUS-CRI (2021) for more details.
Yield Slope (5y - FF)	The slope of the US yield curve calculated as the difference between the US 5-year Treasury rate and the Fed funds rate.
Oil Price	West Texas Intermediate Oil Price. This variable is downloaded from World Bank.
Global Growth Rate	GDP growth rate of G7 economy as a proxy for global growth rate. This variable is downloaded from OECD.
Broad US$	A weighted average of the foreign exchange value of the U.S. dollar against the currencies of a broad group of major U.S. trading partners. This variable is downloaded from FRED, Federal Reserve Bank of St. Louis.

Sources: Data and Corporate Default event for all firm-specific related variables are retrieved from the CRI database, the Credit Research Initiative of the National University of Singapore (NUS CRI), accessed on July 1, 2021.

A.1 Econometrics Tests

In this section, we describe the econometrics tests that we employ to evaluate if the corporate default data in the time rescaled interval follows a time rescaled interval. As pointed out in the main text, our econometrics tests are based on Das et al. (2007), Lando and Nielsen (2010), and Karlis and Xekalaki (2000).

A.1.1 Fisher Dispersion Test

Following Das et al. (2007), Fisher Dispersion Test evaluates if corporate default events in time rescaled interval $\{X = X_i, i = 1, 2, ..., k\}$ is i.i.d, and follows a Poisson distribution with rate c. By fixing the bin size at c, and under the aforementioned stated null hypothesis, we can write the test statistics (W) as:

$$W = \sum_{i=1}^{k} \frac{(X_i - c)^2}{c}$$

W follows a χ^2 distribution with degree of freedom k - 1. In other words, W $\sim \chi^2_{k-1}$. If corporate default events exhibit severe degree of clustering, W will be a large number relative to χ^2_{k-1}. Correspondingly, we will observe a small p-value, and reject the null hypothesis that our data follows a Poisson distribution with rate c.

A.1.2 Potthoff-Whittinghill-Bohning (PWB) Test

As described by Karlis and Xekalaki (2000) and Azizpour et al. (2018), Potthoff-Whittinghill-Bohning (PWB) test evaluates for the moment of the Poisson distribution. In other words, the test assess if the mean of the data is equal to the variance of the data. The test statistics PWB can be written as:

$$PWB = \frac{\sum_{i=1}^{k}(X_i - \bar{X})^2}{\bar{X}\sqrt{2(k-1)}} - \frac{\sqrt{k-1}}{2}$$

where $\bar{X} = \frac{1}{k}\sum_{i=1}^{k} X_i$. According to Karlis and Xekalaki (2000), the asymptotic distribution of the PWB test statistics follows a standard normal distribution. Intuitively, if corporate default events are clustered, and are not able to be sufficiently explained by explanatory variables, we will observe that variance of default data in time rescaling interval to be considerably larger than the mean. In this case, the corresponding p value of this test will be small. This result suggests that the default data in time rescaled interval do not follow a Poisson distribution.

A.1.3 Independence Test

Following Das et al. (2007) and Giesecke and Kim (2011), we use an Autoregressive model to evaluate if default data in time rescaled interval are i.i.d. The Autoregressive model can be written as:

$$X_t = A + BX_{t-1} + \epsilon_t$$

where A and B are parameters to be estimated. Our null hypothesis is that A = c, B = 0, and ϵ_t are i.i.d demeaned Poisson random variables. If our default data in time rescaled interval follows a Poisson distribution with rate c, we should not reject the null hypothesis that A = c, and B = 0.

In contrast, if corporate default events are clustered, we should reject the null hypothesis that A = c, and B = 0. Reject the null hypothesis that A = c imply that the data do not follow a Poisson distribution with rate c, and rejecting the null hypothesis that B= 0 indicates that default events are more persistent than the estimate of our reduced form model. This preliminary finding may indicate evidence of a latent factor or contagion that impact default events.

A.1.4 Kocherlakota-Kocherlakota (KK) Test

As applied by Lando and Nielsen (2010) and Azizpour et al. (2018), the KK test compares the empirical moment generating function of the default data in time rescaled interval with its theoretical counterparts. By first defining the empirical probability generating function as $\phi_k(t) = \frac{1}{k}\sum_{i=1}^{k} t^{X_k}$, we can write the KK test statistics as:

$$KK = \sqrt{n}\frac{\phi_n(t) - exp(\bar{X}(t-1))}{exp(\bar{X}(t^2-1)) - exp(2\bar{X}(t-1))(1+\bar{X}(t-1)^2)}$$

According to Karlis and Xekalaki (2000), the asymptotic distribution of the KK test statistics follows a standard normal distribution. Following Lando and Nielsen (2010) and Karlis and Xekalaki (2000), we select a small t = 0.01. Selecting t along the same range produces a similar result.

A.1.5 Upper Tail Tests

Finally, we use two different variants of upper tail tests to evaluate the tail properties of the default data in time rescaled interval. Following Das et al. (2007), we would like to assess if the upper quartile of the empirical default in time rescaled interval with rate c, is explicitly larger than the upper quartile of the theoretical poisson distribution with rate c.

The first variant of the upper quartile test can be applied by select a bin size of c, and suppose there are k observations in this data sample. Define M as the sample mean of the upper quartile of empirical distribution of $\{X = X_i, i = 1, 2, ..., k\}$. We simulate 10,000 data sets. Each dataset consists of k random variables that are i.i.d, which follows a Poisson distribution with rate c.

As suggested by Das et al. (2007), we can calculate the p-value as the fraction of the simulated data sets where the sample upper-quartile size (mean) exceeds the actual sample mean, M. Correspondingly, we can construct the second variant of the upper quartile test by replacing the mean with the median.

Table A.2: Out-Of-Sample Default Risk Assessment: Area Under the Curve (AUC)

Regions	U.S.	Europe	Asia Pacific	Emerging Markets
Augmented Model	0.968	0.921	0.929	0.827
Original Model (Das et al. 2007)	0.964	0.917	0.929	0.777

This table presents the Out-of-Sample Area Under the Curve (AUC) measure for different Proportional Hazard models, across different regions.