FEDERAL RESERVE BANK OF BOSTON™

RISK AND POLICY ANALYSIS UNIT

Working Paper
RPA 17-01
May 8, 2017

Estimating Loss Given Default from CDS under Weak Identification

Lily Y. Liu

Risk and Policy Analysis (RPA) Working Papers present economic, financial and policy-related research conducted by staff in the Federal Reserve Bank of Boston's Risk and Policy Analysis Unit. RPA Working Papers can be downloaded without charge at: http://www.bostonfed.org/bankinfo/qau/wp/

Estimating Loss Given Default from CDS under Weak Identification

Lily Y. Liu *
Federal Reserve Bank of Boston

November 1, 2016

Abstract

This paper combines a term structure model of credit default swaps (CDS) with weak-identification robust methods to jointly estimate the probability of default and the loss given default of the underlying firm. The model is not globally identified because it forgoes parametric time series restrictions that have aided identification in previous studies, but that are also difficult to verify in the data. The empirical results show that informative (small) confidence sets for loss given default are estimated for half of the firm-months in the sample, and most of these are much lower than and do not include the conventional value of 0.60. This also implies that risk-neutral default probabilities, and hence risk premia on default probabilities, are underestimated when loss given default is exogenously fixed at the conventional value instead of estimated from the data.

JEL Codes: G12, G13, C58, C13, C14

*I am grateful to my Ph.D. advisor, Andrew Patton, for his guidance and support. I also thank Tim Bollerslev, Federico Bugni, Shakeeb Khan, Jia Li, George Tauchen, and seminar participants of the Duke Financial Econometrics Workshop and the Federal Reserve Bank of Boston for their insightful comments, and Daniel Smith for excellent research assistance. All remaining mistakes are mine. The views expressed in this paper are my own and do not necessarily reflect those of the Federal Reserve Bank of Boston or Federal Reserve System. Contact e-mail: lily.liu@bos.frb.org.

1 Introduction

Since its inception in the mid-1990s, the credit default swap (CDS) market has seen incredible growth, with notional outstanding reaching tens of trillions of dollars by 2005.[1] Correspondingly, there has been a growing interest in measuring and understanding the risk-neutral credit risk reflected in CDS prices. This credit risk can be decomposed into two fundamental components: the risk-neutral probability of default (PD) and the risk-neutral loss of asset value given occurrence of a default event (LGD). However, their joint estimation is complicated because these two components contribute to CDS prices (S) in an approximately multiplicative manner, i.e., $S \approx LGD \times PD$. To circumvent this identification issue, the traditional CDS pricing literature fixes loss given default at an exogenous value and focuses on estimating default probabilities. For U.S. corporates, LGD is usually set around 0.60, a value obtained from historical data on observed defaults.

While this simplifying assumption is benign for certain applications, such as fitting CDS spreads,[2] there are important financial applications that require separate estimates of one or both of these components. Examples include studying the risk premium associated with either component, or valuing or hedging related credit-sensitive assets whose payoffs are affected by PD or LGD differently than CDS payoffs.[3] Even if probability of default is the sole object of interest, fixing LGD incorrectly will lead to distorted estimates. In response to the need for unbiased estimates, a literature on joint estimation has emerged. The common identification strategy in these papers is to use multiple defaultable assets written on the same underlying firm. These assets share a common probability of default (and possibly common LGD), but PD and LGD affect their prices differently (due to contract differences). Thus, harnessing the information in the cross-section of prices can allow for joint estimation.[4]

This paper adds to this literature, pairing a CDS term structure model with weak-identification robust econometric methods. The model and inference methods are both new to the joint estimation literature, and their combination allows LGD and PD to be estimated without relying on parametric time series restrictions that are difficult to verify in the data. In addition, by employing the term structure of CDS as the multiple assets for

[1] BIS Semiannual OTC derivatives statistics, starting from the May 2005 issue, accessible at http://www.bis.org/publ/otc_hy1405.htm

[2] Houweling and Vorst (2005) show that many fixed values of LGD yield similar results for fitting CDS spreads.

[3] This is especially the case for related credit derivatives such as digital CDS, junior debt instruments and recovery swaps.

[4] Pan and Singleton (2008) is an early paper that adopts this identification strategy applied to the term structure of CDS, see paper for discussion.

joint identification rather than combining CDS with equity options or junior ranked debt as in some papers, lack of cross-market integration is not a concern, and data is available for a larger cross-section of firms.[5]

This paper has two main objectives. Firstly, I jointly estimate LGD and default probabilities without restrictive parametric assumptions and without requiring additional data beyond multiple maturities of CDS. As a result of imposing fewer structural assumptions, this model is not globally identified. Thus, I employ robust econometric methods that allow for valid inference regardless of the strength of model identification. Secondly, I estimate this model and obtain confidence sets for LGD for a selection of U.S. investment grade and high yield firms. I then examine how the estimates of LGD under the cross sectional model compare to the conventional level of 0.60. My results show that for almost half of the firm-months, LGD is precisely estimated, i.e. confidence sets are small, and the estimates are approximately between 0.05–0.20. Furthermore, when LGD is precisely-estimated, the value of 0.60 is almost always rejected. As a direct consequence, risk neutral default probability is underestimated using conventional methods, which also implies that risk premia associated with default probability are underestimated in the existing literature.

This paper differs from existing work in three main ways. Firstly, I directly model risk-neutral expected LGD and PD term structures at a point in time. In place of time series restrictions, I assume that CDS spreads over short periods of time (one calendar month in the base case) are generated from the same model, so the model can be estimated independently each month. In contrast, most of the joint estimation literature augments the "reduced-form intensity model" framework of the traditional CDS pricing literature (see Duffie (1998) and Duffie and Singleton (1999)) to allow for stochastic LGD.[6] In these models, the default event is defined as the first jump of a Poisson process with stochastic intensity, and thus the models consist of parametric specifications for the dynamics of the true (latent) default intensity process and the price of risk (e.g., Madan et al. (2006), Pan

[5] Conrad et al. (2013) pair equity options with CDS and occasionally find negative estimates of LGD, which they attribute to differences in price discovery between CDS markets and equity option markets. Further, equity option data is only available and reliable for larger publicly traded firms. Schläfer and Uhrig-Homburg (2014) pair senior CDS contracts (which are readily available) with junior subordinated CDS and LCDS for which there is limited data.

[6] Das and Hanouna (2009) and Schläfer and Uhrig-Homburg (2014) are not reduced-form intensity models. Das and Hanouna (2009) is the paper whose modelling framework is most similar to ours, as they also aim to extract "point-in-time" risk-neutral expectations about credit risk. However, they use a calibration (in contrast to econometric) approach and fit a dynamic jump-to-default model with state dependent default intensity. Schläfer and Uhrig-Homburg (2014) do not use a time series model, but rather use the ratio of senior to junior CDS prices to identify unconditional moments of the risk-neutral distribution of LGD, which they model using the beta distribution.

and Singleton (2008), among others[7]). The term structure of risk neutral LGD and PD, and other objects,[8] can then be computed from these two central components. However, the richness of these models comes at a cost. The parametric assumptions on default dynamics are difficult to verify, and there is no consensus on which of the numerous model specifications is best. Further, it is uncertain how sensitive estimates are to model specification. Empirical results from different studies are difficult to compare as they do not generally use the same price data, sample period, or reference entities. By employing a minimally parameterized model, this paper provides estimates of risk neutral loss given default robust to the default intensity specification.

Secondly, the term structure of loss given default, which is assumed to be flat in the base case, is estimated less restrictively than in the existing literature. In Section 6.2, the term structure of LGD is allowed to be linear, and implications on joint identification are investigated. Even though LGD term structure is constant over each estimation period, the model is estimated independently each month, so I obtain a time series of LGD estimates. This is an improvement over existing literature in which the LGD is modelled as a constant over the entire multi-year sample period as in Pan and Singleton (2008), Elkamhi et al. (2010), and Schneider et al. (2011). A few papers do estimate time-varying LGD, but require it to be a direct function of the default probability. For example, LGD is modelled as an exponential affine function with positive correlation to default intensity in Madan et al. (2006), and as a linear probit in Das and Hanouna (2009). In this model, no functional relationship between LGD and PD is specified.

Finally, this paper provides a novel application of weak-identification robust methods as the first paper to employ such methods towards jointly estimating LGD and PD. Existing joint estimation papers have worked around the potential identification issue by using parametric time series models for a cross-section of defaultable assets, that are then estimated after assuming strong identification. The only papers to investigate and offer evidence of model identification are Pan and Singleton (2008) and Christensen (2005) using simulation methods, and Christensen (2007) using actual CDS data. By using the robust econometric methods in Stock and Wright (2000), I can relax the parametric time series assumptions, and only impose shape restrictions on the term structures of LGD and PD.

Weak identification arises when models are strongly identified for most of the parameter space, but not identified for a particular region of the parameter space; when model

[7] Also, Le (2014), Song (2007), Christensen (2005) , Christensen (2007), Elkamhi et al. (2010), and Schneider et al. (2011)

[8] In addition, the objective default probabilities and various risk premia can be computed. The time series evolution of all these objects can also be studied in this framework.

parameters are local to the region of non-identification, the model is said to be weakly identified. Within the broad weak identification literature, a large portion of applications and theoretical work, including Stock and Wright (2000), deal with the weak instrumental variables regression setting.[9] Andrews and Cheng (2012) covers inference under weak identification for a large complementary set of models (generally distinct from the weak IV setting), whose criterion function depends on a parameter that determines the strength of identification.[10]

The model I use in this paper to jointly estimate LGD and PD does not directly fit in the weak IV setup nor in the family of models considered in Andrews and Cheng (2012). In this model, a general criterion function does depend on a parameter that determines strength of identification, as in Andrews and Cheng (2012), however, when that parameter is in the region of non-identification, which occurs when the PD term structure is flat, the model is not completely non-identified, but is rather set identified, or partially identified.[11]

The outline of the paper is as follows. Section 2 introduces the CDS data. Sections 3 and 4 describe the model and estimation methodology. In Section 5, I present and analyze the estimated LGD confidence intervals and elaborate on the main findings. Section 6 discusses robustness of the results and model extensions, and Section 7 concludes.

2 Credit Default Swaps and Spread Data

After a brief description of credit default swap contracts, this section presents an overview of the Markit CDS data used in this study. Then, I introduce the cross section of CDS reference entities selected for this study, and present firm characteristics and summary statistics for the CDS prices.

[9] Some empirical applications in macroeconomics and macro-finance include estimation of the coefficient of risk aversion for CRRA utility, which is weakly identified in the Euler equation in the consumption-CAPM model, see Stock and Wright (2000), and estimation of the New Keynesian Phillips Curve, see Canova and Sala (2009) and Nason and Smith (2008).

[10] Some examples in the Andrews and Cheng (2012) framework include nonlinear regression with a multiplicative parameter and estimation of ARMA(1,1), which is not identified when the autoregressive and moving average coefficients are equal.

[11] A non-linear function of the "level" of the PD term structure and the "level" of LGD term structure is identified, but these two objects are not separately identified. Thus the model parameters are set or partially identified.

2.1 CDS spread data description

A credit default swap is an over-the-counter derivative written on a risky reference bond that allows for the transfer of the bond's default risk between two parties for an agreed on length of time. The CDS buyer pays a periodic premium (quarterly, for corporate contracts) to the CDS seller in exchange for the seller guaranteeing the value of the bond after a default event. If a default event occurs during the contract lifetime, premium payments stop (accrual payments are accounted for), and the CDS seller will compensate the loss of bond value due to default.[12]

CDS prices used in this paper are composite quotes from Markit Group. Markit collects CDS quotes from individual dealers, filters out unreliable prices, performs mark-to-market adjustments, and aggregates them into a daily composite quote for the following maturity points: 6-months, 1–5, 7, 10, 15, 20, and 30 years. See Markit Group Ltd. (2010) for details. Since 15-year and above contracts are not as actively traded, I only use the 8 CDS contracts with maturity points 10 years or less, effectively limiting estimation of the forward default and LGD curves to up to 10 years as well.

The sample period spans January 2004 to February 2016, for a total of 146 months. In October 2005, Markit begins reporting 4-year CDS spreads, and I add this maturity point to the study. In April 2009, the CDS Big Bang implemented changes for CDS contracts and the way they are traded. Neither of these changes affect estimation since the model is estimated independently each month, however, when looking at the estimation results, I check whether there are any systematic differences before and after either date, and I do not find any large differences (see Section 5.1).

I choose a collection of 30 U.S. corporate issuers (listed in table A.1 in the appendix), that span a variety of industries and credit ratings. Twenty firms are chosen from the CDX North American Investment Grade CDS index CDX.NA.IG series 17, and 10 firms from the North American High Yield CDS index CDX.NA.HY Series 17. These indices are issued every 6 months and collect the most liquid single-name entities from their respective credit class at the time; Series 17 was issued in September 2011.[13] I randomly

[12] The majority of CDS contracts in the market are unbacked, meaning that the CDS buyer does not hold the actual defaultable bond. The loss amount that the CDS seller is responsible for is determined by auction price of the defaulted bonds. The auction is overseen by the ISDA and usually takes place a few months after the default event. See Markit Group Ltd. (2010) or Barclays Capital (2010) for additional details. Since the CDS Big Bang in April 2009, the CDS market has moved towards different pricing conventions, so that there is now upfront payment and reduced coupon, however the format of Markit quotes is unchanged.

[13] CDX.NA.IG contains single name CDS from 125 firms, and CDS.NA.HY contains single name CDS from 100 firms. Investment grade firms have long-term credit ratings from AAA/AAa (highest) to BBB/Baa2 (lowest). Firms with credit rating BBB-/Baa3 and lower are considered high yield ("speculative" or "junk" are other common names).

selected issuers covering each sector listed in CDX, after excluding issuers with limited data availability in the earlier years. If, for a given firm-month, there exist at least 18 days with at least 7 CDS prices per day, then that firm-month is included in the sample, otherwise, it is dropped. Across all issuers, 166 firm-months were dropped, though this includes 72 firm-months, from the beginning or end of the sample period, that did not have any reported CDS data, in addition to 94 firm-months with sparse CDS prices. [14]

This study is limited to XR (no restructuring) contracts on senior unsecured bonds traded in US dollars. XR contracts were adopted as the conventional contract for U.S. corporates after the CDS Big Bang in 2009, and thus are more commonly traded than contracts with other restructuring clauses. However, prior to the Big Bang, MR (modified restructuring) contracts were more popular.[15]

Credit spreads in the form of corporate yield spreads can be constructed from defaultable bond data and a reference risk-free rate; however Longstaff et al. (2005) show that corporate yield spreads are on average 1.2-2 times higher than CDS spreads, depending on credit rating, and they attribute the extra spread mainly to illiquidity effects. Thus, CDS spreads are favored over corporate bond spreads for studying default risk. Certainly, CDS prices are not immune to liquidity risk themselves. Liquidity premium in CDS prices has been studied, but there is no consensus on its size, or whether the CDS seller or lender receives the premium. Other risk factors (unrelated to issuer default) that may affect CDS prices are further discussed in Section 3.1.

2.2 Summary Statistics for CDS spreads

Table 1 lists the CDS reference entities and presents summary statistics for CDS spreads from 1, 5, and 10-year contracts.[16] The median 5-year spread ranges from 36 bps (Union Pacific Corp.) to 571 bps (Advanced Micro Devices) while the mean 5-year spread ranges from 39 bps (Union Pacific Corp.) to 1191 bps (Radioshack). In this sample, the high yield contracts have average spreads that are around 3–5 times larger than average investment grade spreads of the same maturity, and similarly, the average HY spread standard deviation per issuer is around 4 times higher than average IG standard deviation.

[14] Valero Energy Corp (VLO) did not have reported CDS prices until September 2005. In addition, the CDS data series for L Brands Inc. (LTD), Radioshack Corp.(RSH), and General Electric Capital Corp (GEcc). ended in Mar. 2013, Jan. 2015, and Jan. 2016, respectively. Radioshack Corp. experienced a bankruptcy credit event around Jan. 2015; the ISDA Event Determination Date is listed as Feb. 6, 2015.

[15] Berndt et al. (2007) study 5-yr corporate CDS from 1999–2005 and find that the difference between MR and XR contract prices is very small for high quality firms, but increases as the level of CDS prices increases. They also estimate that on average (over 2000 firms), MR contract prices are 6–8% higher than XR prices.

[16] Additional information on the full names, credit ratings and industry sector of the firms can be found in Table A.1 in the Appendix.

Table 1: **Summary Statistics of CDS spreads**

	1 year		5 year			10 year		5 year			
	mean	sd	mean	med	sd	mean	sd	skew	ac(1)	ac(10)	ac(20)
					Investment Grade						
GEcc	85	163	117	70	133	128	112	2.53	1.00	0.96	0.92
SWY	35	24	142	92	107	190	140	1.00	1.00	0.99	0.97
WHR	50	85	117	95	87	151	79	1.94	1.00	0.97	0.93
UNP	15	17	39	36	19	59	15	1.75	1.00	0.93	0.87
UNH	37	57	76	45	69	98	64	1.66	1.00	0.98	0.95
MO	38	40	78	71	42	108	42	0.80	1.00	0.97	0.93
CSX	22	28	52	45	32	73	28	2.11	1.00	0.96	0.93
CSC	34	54	108	88	85	151	96	1.95	1.00	0.97	0.94
AZO	22	23	64	62	25	92	22	1.33	1.00	0.93	0.85
AXP	63	129	83	49	97	92	78	3.19	1.00	0.95	0.88
ARW	32	31	100	87	43	141	49	0.89	1.00	0.95	0.88
AA	85	162	193	170	168	231	166	1.37	1.00	0.97	0.94
YUM	24	23	72	65	38	99	46	2.50	1.00	0.96	0.92
COP	17	19	46	39	40	64	44	4.92	1.00	0.96	0.92
VLOC	50	62	124	117	68	158	67	0.53	1.00	0.96	0.93
COX	23	27	70	65	31	102	31	2.00	0.99	0.90	0.79
CAG	16	11	57	55	25	85	37	0.73	1.00	0.96	0.92
AEP	18	15	46	41	19	66	21	0.50	1.00	0.95	0.90
APC	55	93	106	84	96	135	95	2.64	1.00	0.93	0.85
XRX	71	92	160	136	88	203	79	1.71	1.00	0.95	0.90
					High Yield						
LTD	96	143	176	178	133	206	119	1.54	1.00	0.97	0.93
MGIC	546	719	530	338	529	485	420	1.04	1.00	0.98	0.95
AMD	429	715	740	571	625	761	517	2.31	1.00	0.97	0.95
GT	215	271	451	429	225	497	174	1.53	1.00	0.95	0.88
TSG	311	642	539	368	617	557	524	2.94	1.00	0.98	0.95
RSH	1326	4369	1191	197	3275	859	1716	6.00	1.00	0.95	0.91
RCL	183	358	317	202	290	348	226	2.64	1.00	0.97	0.93
THC	200	181	505	458	204	551	160	1.94	1.00	0.96	0.92
WY	41	47	107	86	63	143	60	0.61	1.00	0.97	0.94
CMS	77	79	162	157	87	188	76	0.60	1.00	0.97	0.95

Notes: This table presents the sample means and standard deviations of 1, 5, and 10-year daily CDS spreads for all reference entities under study for the sample period of January 2004 - February 2016. Also, the sample median, skewness and 1, 10, and 20-lag autocorrelations of the 5-year spreads are reported. The reference entities are identified by their Markit ticker (COX and GEcc are abbreviated Markit tickers). The full firms names can be found in appendix table A.1.

The average CDS spread term structure is upward-sloping, and generally we only see inverted term structures in times of credit distress (as with yield curves). In addition, CDS spreads are right skewed and highly serially correlated. Table 1 presents the skewness of 5-year spreads, and the autocorrelation estimates for daily 5-year spreads for 1, 10, and 20-lags. The 5 year CDS spread skewness is under 3.0 for 27 out of 30 firms, the major outlier being Radioshack, which experienced a bankruptcy during the sample period, and had steadily and sharply increasing CDS spreads leading up to the credit event. The 20-lag autocorrelations range from 0.79 (COX) to 0.97 (SWY). These results suggest that CDS spreads are highly persistant, and potentially close to a unit root, however, this does not pose a problem for the estimation procedure because only the estimating equations (for minimum distance estimation, see Section 4.1) are required to be covariance stationary.

Overall, CDS spreads across firms are moderately correlated as the average pairwise correlation of 5-year spreads across firms is 0.47. However, there are substantial differences among firm-pairs, as these pairwise correlations range from -0.50 to 0.93. Figure 1 plots the time series of 1, 5 and 10-year spreads for four representative firms. For many firms in our sample, like General Electric Capital Corporation (GEcc) and CSX Corporation (CSX), CDS spreads are very low between 2004 and 2007, peak during the financial crisis, and then fall to levels a little higher than spreads in the pre-crisis era. In contrast, there are also several firms, for example Altria (MO) and Safeway (SWY), whose CDS spreads exhibit pronounced peaks outside of the financial crisis. It is possible that the sample is biased towards firms with higher credit risk in the middle period since they were chosen from CDX Series 17 indices, which are composed of the most active single names around September 2011. However, 19 and 5 of the 30 firms were also listed in Series 1 CDX.NA.IG and CDX.NA.HY, respectively, which were on-the-run at the beginning of the sample period (October 2003 - March 2004), and 17 and 7 of the firms are listed in the Series 25 CDX.NA.IG and CDX.NA.HY, respectively, which were on-the-run at the end of the sample period (October 2015–March 2016).

3 CDS Term Structure Model

In this section, I describe the CDS pricing framework used in this paper. I introduce the term structure model for the forward default probability curve and LGD curve, and show that the model is strongly identified for most but not all of the parameter space, so that econometric methods robust to weak identification are necessary.

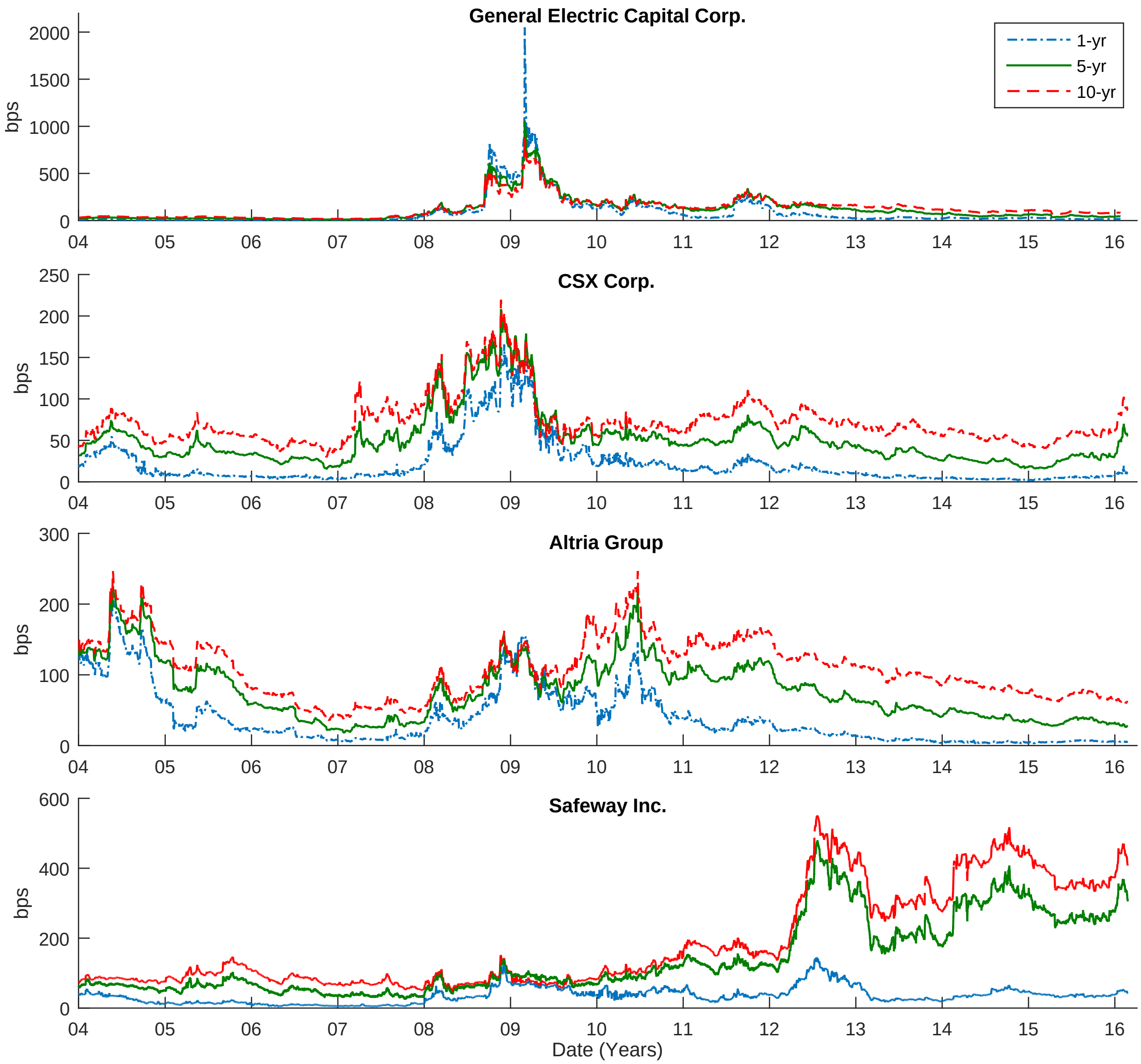


Figure 1: *This figure plots the time series of 1-year, 5-year and 10-year CDS spreads for four representative firms. The sample period runs from January 2004 to February 2016, and the firms plotted are General Electric Capital Corporation (GEcc), CSX Corporation (CSX), Altria Group (MO) and Safeway Inc. (SWY).*

3.1 A discrete-time framework for default

I model the CDS market at time t using a discrete-time model similar to Conrad et al. (2013), but allowing for more general (non-flat) term structures for PD and LGD, as described below.

I assume that at time t, CDS contracts with maturities of n quarters are struck, for $n = \{2, 4, 8, 12, 16, 20, 28, 40\}$. The CDS spread $s_t^{(n)}$ is the annualized premium paid to insure \$1 of an underlying corporate bond over the life of the n-quarter CDS contract, and payment is made at the end of each quarter that the underlying entity does not default, for quarters $j = 1, ..., n$. The time-t risk-neutral expectation of the probability that the underlying bond will default j quarters from time t, conditional on survival through the $j-1$-th quarter, is represented by $q_{j,t}$. The set $\{q_{j,t} : j = 1, ..., 40\}$ is referred to as the forward probability of default term structure at time t, or simply PD curve. Payments are discounted using a zero coupon term structure, which is taken as known and extracted from data on the U.S. treasury yield curve. The j-quarter discount rate at time t (a cumulative spot rate, not a forward rate) is denoted by $\bar{d}_{j,t}$. The term structure of loss given default, also called the LGD curve, is given by $\{L_{j,t} : j = 1, ...40\}$, where $L_{j,t}$ is defined as the time-t risk-neutral expectation of the proportional loss of face value of the underlying bond, given default j quarters from time t. Note that this definition is the commonly used loss convention *fractional recovery of face value* (or par value).[17] Guha and Sbuelz (2005) provide empirical evidence supporting this loss convention, and it is the most natural choice given CDS contract wording.[18]

Under the model assumptions above, the present value of the CDS premium leg is

$$\sum_{j=1}^{n} \prod_{k=1}^{j-1} (1 - q_{k,t}) \left\{ (1 - q_{j,t}) \bar{d}_{j,t} s^{(n)} + q_{j,t} \bar{d}_{j,t} \times \frac{1}{2} \times \frac{s_t^{(n)}}{4} \right\} \tag{1}$$

and the present value of the CDS protection leg is

$$\sum_{j=1}^{n} \prod_{k=1}^{j-1} (1 - q_{k,t}) \left\{ (1 - q_{j,t}) \bar{d}_{j,t} \times 0 + q_{j,t} \bar{d}_{j,t} L_{j,t} \right\}. \tag{2}$$

Equating these two payoff legs, we can solve for the no arbitrage spread of an n-

[17] The other main loss characterizations in the literature are fractional recovery of market value and fractional recovery of treasury. See Duffie and Singleton (1999) and Madan et al. (2006) for comparisons of the fractional recovery of market value and fractional recovery of treasury assumptions for LGD.

[18] Guha and Sbuelz (2005) show that after a default event, bonds of the same seniority are observed to recover the same proportion of bond face value, irrespective of maturity. This empirical fact is generally only consistent with the fractional loss of face value framework.

quarter CDS contract struck at time t

$$s_t^{(n)} = \frac{4 \times \sum_{j=1}^{n} \bar{d}_{j,t} \prod_{k=1}^{j-1}(1-q_{k,t}) q_{j,t} L_{j,t}}{\sum_{j=1}^{n} \bar{d}_{j,t} \prod_{k=1}^{j-1}(1-q_{k,t})(1-\frac{1}{2}q_{j,t})}. \tag{3}$$

In the base case, I further simplify the model by assuming that the term structure of LGD is flat, i.e. $L_{j,t} = L_t \; \forall j$. Later in section 6.2, I consider a linear specification for the term structure of LGD, and explore the implications on joint identification.

As mentioned in the introduction, LGD in this model (including both the flat and linear term structures) is modelled less restrictively than in the existing literature. Most intensity models estimate LGD as a constant over a sample period of multiple years, as in Pan and Singleton (2008), Schneider et al. (2011), and Elkamhi et al. (2010). In a small number of papers, LGD is time-varying, but to keep the model tractable, LGD is modelled as a function of PD. For example, Madan et al. (2006) restrict LGD to be exponential affine in and positively correlated with default intensity.[19] In this paper, LGD is estimated independently each month, so even though the model lacks dynamic features, we still obtain a sequence of monthly estimates of LGD. Additionally, there are no functional restrictions between LGD and PD. The obvious drawback is that we lose any efficiency gains that would be achieved if the relationship between LGD and PD were correctly specified.

Under the flat term-structure assumption for LGD, the CDS premium expression is simplified to

$$s_t^{(n)} = L \frac{4 \times \sum_{j=1}^{n} \bar{d}_{j,t} \prod_{k=1}^{j-1}(1-q_{k,t}) q_{j,t}}{\sum_{j=1}^{n} \bar{d}_{j,t} \prod_{k=1}^{j-1}(1-q_{k,t})(1-\frac{1}{2}q_{j,t})}. \tag{4}$$

In this pricing equation (4), there are $n+1$ unknowns: $q_{j,t}$ for $j = 1, \ldots, n$, which map out quarterly points on the forward PD curve, and L_t, the expected loss given default. In my analysis, there are $n = 40$ quarters in total, so to make estimation feasible, I reduce the dimensionality of the model by adopting a family of flexibly-shaped curves f for the default probability term-structure, i.e. $q_{j,t} = f(j; \beta_t)$.

[19] Most intensity models in the joint estimation literature specify LGD restrictively, and likely do so in order to stay within the class of affine term structure models. Given a general affine specification (in state variables) for default intensity, LGD is required to be either a constant or exponential affine in the default intensity process. Exceptions are Christensen (2005) and Christensen (2007), where both default intensity and LGD are affine in state variables, which leads to a quadratic term structure model for CDS prices. The quadratic term structure model is the highest order model for which closed form solutions exist. Elkamhi et al. (2010) also use a quadratic term structure model for CDS prices, but they allow default intensity to be quadratic in latent factors, and then must model LGD as a constant to ensure close-form solutions for CDS prices.

3.2 A factor model for the term structure of default probabilities

Determining a suitable model for the forward default curve is difficult since the default curve is not observed, even *ex post*. However, I can construct a set of "proxy" PD curves for the unobserved PD term structures, conditional on a flat term structure for LGD (as assumed in this model). Then I can find a factor model that captures most of the variation of the proxy curves.

Since construction of the proxy curves requires a fixed value of LGD, yet the resulting curves may be sensitive to the particular choice, I consider various levels for the LGD term structure, {0.1, 0.2, 0.4, 0.6, 0.8, 1}, and construct a set of proxy curves for each value of LGD. So, for each value of LGD, assuming a flat default probability between consecutive CDS maturities, I "strip" each daily CDS price curve from lowest to highest maturity to get a step function with 8 steps as an approximate forward PD curve. See Figure A.1 in the appendix for an example for CDS data from one day.[20] Then, for each fixed value of LGD, I conduct principle components analysis on the set of 8-dimensional vectors that represent daily implied forward default probability curves. The analysis is conducted separately for each of the 30 firms and on the pooled set of implied PD curves (after demeaning for each issuer) across all assets. Figure 2 presents the first three principal components for the pooled set of PD curves for all six values of LGD. For each fixed value of LGD, 95% to 98% of all variation in the pooled PD curves is captured by the first three components. Further, the components visually resemble level, slope and curvature loadings. These results suggest that using a 3-factor "level-slope-curvature" model for the forward default probability curve is a good approximation.

[20]The 8-step default probability curve assumes constant forward default probability per quarter over the following time intervals (units in years) whose endpoints correspond to CDS contract maturities: $(t, t+0.5]$, $(t+0.5, t+1]$, $(t+1, t+2]$, $(t+2, t+3]$, $(t+3, t+4]$, $(t+4, t+5]$, $(t+5, t+7]$, $(t+7, t+10]$. Resulting proxy PD curves with values not in the $(0,1]$ interval are discarded from the PCA.

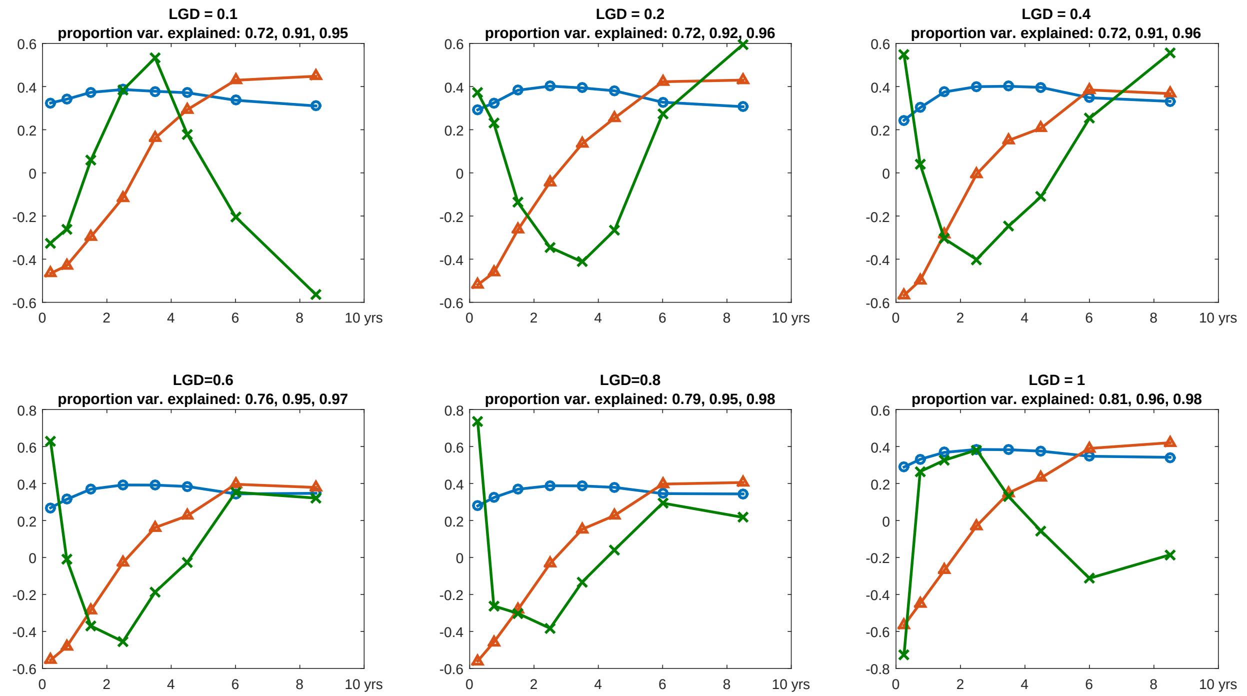


Figure 2: *This figure plots the first three principal components of the set of pooled (over reference entities) daily proxy default probability curves implied for six values of LGD. The set of proxy PD curves are demeaned per reference entity prior to analysis.*

3.3 Nelson-Siegel curves: a level-slope-curvature model

Based on the results of the principal component analysis, I propose modelling the forward default probability term structure $q_{j,t}$ using Nelson-Siegel curves, which are a family of curves composed of a linear combination of three components resembling level, slope, and curvature.

Equation (5) presents the form of the Nelson-Siegel curve used in this paper, which is the Diebold and Li (2006) reparameterization of the original form that Nelson and Siegel (1987) propose to fit U.S. government yield curves. The three β coefficients determine the weights of each component, while the fourth parameter λ determines the rate of decay of the exponentials in the function, which directly relates to the shape of the slope component and the location of the "hump" in the curvature component. See Figure 3 for an illustration of the three Nelson-Siegel curve components.

$$q_{j,t} = f(j;\boldsymbol{\beta}_t) = \beta_{1t} + \beta_{2t}\frac{1-e^{-\lambda j}}{\lambda j} + \beta_{3t}\left(\frac{1-e^{-\lambda j}}{\lambda j} - e^{-\lambda j}\right) \tag{5}$$

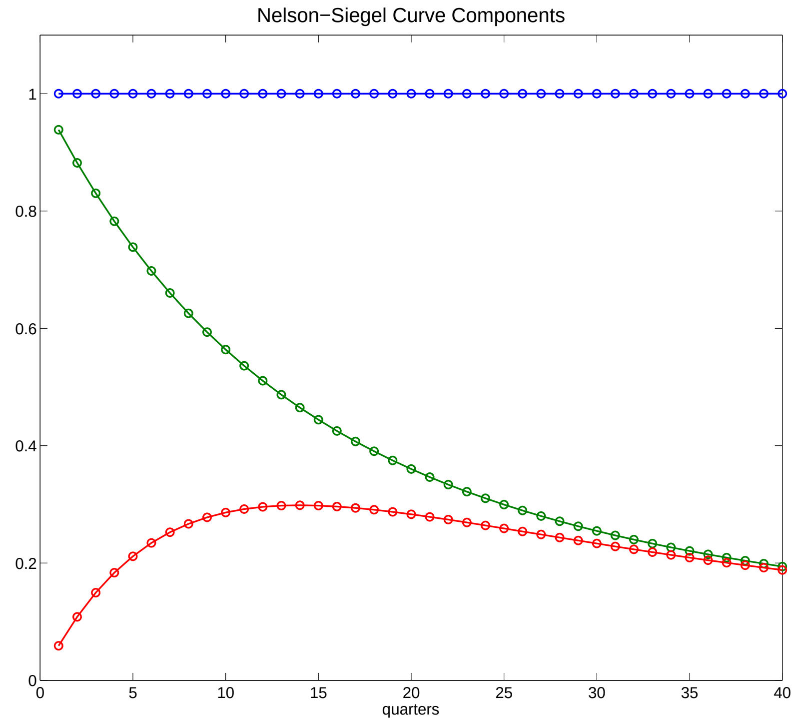


Figure 3: *This figure presents the three components of the Nelson-Siegel curve with λ fixed so the hump of the third component is at 3.5 years (42 months).*

Nelson-Siegel curves have been used extensively in research and in practice, including by central banks.[21] They are noted for their parsimony and ability to match the many shapes of observed yield curves, including flat, upward or downward sloping with varying convexity, humped, and mildly S-shaped.

As mentioned above, the λ parameter in the Nelson-Siegel curve determines the exact shape of the slope and curvature components. For larger values of λ, the slope component decays slower and the curvature component reaches its maximum later. In the yield curve literature, λ is often fixed to simplify yield curve estimation to OLS. Nonlinear methods can be used to estimate a free λ; Nelson and Siegel (1987) employ a grid search over λ, but the estimates are unstable over time, and Annaert et al. (2013) caution that for certain values of λ, there is a high degree of collinearity among the 3 components.

In this application, the CDS pricing equation is already nonlinear, but to avoid adding an additional layer of nonlinearities, I fix λ guided by the results of the principal components analysis. In the pooled PCA results, the estimated third principle component ("curvature") has a pronounced peak at the 4th step or 5th step, which is centered at the 2.5 or 3.5 year maturity, respectively (see Figure 2). Therefore, I fix $\lambda = 0.1281$ so the hump of the third component is at 3.5 years. In Section 6.1, I also estimate the model using other

[21] See Annaert et al. (2013) for discussion.

values of λ as a robustness check, including the value of λ such that the hump of the third Nelson-Siegel component is at 2.5 years. The main conclusions of my analysis are unchanged. This value of λ is somewhat similar to the value used for yield curve modelling in Diebold and Li (2006), who choose λ to locate the maximum of the curvature component at 2.5 years because 2–3 year maturities are typical for humps and troughs in yield curves in the literature.

This model abstracts from risk factors (unrelated to issuer default) such as illiquidity (as mentioned in section 2.1, counterparty and contagion risk. Counterparty risk is the risk to each party of the contract that the other will not be able to fulfill their contractual obligations. Contagion (see Bai et al. (2015)) is the risk in credit markets, distinct from default event risk, that default of systemically important firms will be contemporaneous with a drop in the market portfolio. Bai et al. (2015) show that all but a few basis points of credit spreads that attributed to the risk premium on default probability in standard (doubly stochastic) intensity models is actually contagion risk premium. In this model, these other risk premia are subsumed into the estimates for risk neutral default probability and LGD. This is the standard approach in the joint estimation literature because accounting for these two other effects can be very complicated, and generally requires model extensions and additional data. It is difficult to say how illiquidity and counterparty risk premia will be divided between risk neutral LGD or risk neutral probability of default. However, contagion risk will likely be soaked up in the risk neutral default probability estimates because in accordance with theoretical contagion models, only the likelihood of default probability is directly affected through contagion channels: the probability that a given firms will default increases when a systematically important firm defaults, and there is no direct impact on recovery rates.

3.4 Joint Identification of Default Probabilities and Loss Given Default

As suggested in Pan and Singleton (2008), when LGD is modelled as the fractional loss of bond face value given default, identification of both probability of default and loss given default can possibly be achieved by exploiting both short and long term CDS contracts because their prices are affected differently by the two components. Pan and Singleton investigate and confirm the effectiveness of this identification strategy using simulated data under a model with log-normal default intensity and constant LGD. However, in the model used in the analysis below, it is easy to see that there is a subset in the parameter space in which the model is not identified. When the forward probability of default has a flat term structure, i.e. $q_{j,t} = q_t$, then the expression for the CDS price reduces to:

$$s_t^{(n)} = L_t \frac{4 \times \sum_{j=1}^{n} \bar{d}_{j,t} \prod_{k=1}^{j-1} (1-q_t) q_t}{\sum_{j=1}^{n} \bar{d}_{j,t} \prod_{k=1}^{j-1} (1-q_t)(1-\frac{1}{2}q_t)} = L_t \frac{q_t}{1-\frac{1}{2}q_t} \text{ for } n = 2, 4, 8, 12, 16, 20, 28, 40\,; \quad (6)$$

and it is apparent that L_t and q_t are not point identified. However, as mentioned in the introduction, it is interesting to point out that the identification issue here is somewhat different than that of most applications in the weak-identification literature because, even in the problematic region in the parameter space, the *nonlinear function of* L_t *and* q_t on the right-hand side of Equation (6) is identified, which restricts the values of L_t and q_t to a set that is smaller than the logical range from 0 and 1 for each, so the model is set (or partially) identified.[22] Furthermore, when the PD curve is close to this non-identified region, the model is *weakly identified:* a criterion function such as standard GMM or nonlinear least squares is relatively flat with respect to L_t and β_{1t}, so standard asymptotics (and standard t and QLR tests) do not provide good approximations and are thus invalid.

Under the Nelson-Siegel parameterization, the non-identified region (or rather, set-identified region) is characterized by $\beta_{2t} = \beta_{3t} = 0$, and since then $q_t = \beta_{1t}$, the subvector (L_t, β_{1t}) is not point-identified.

Also, it is worth noting that this weakly identified region nests the case when default probabilities are very low, i.e. β_{1t}, β_{2t} and β_{3t} are close to 0, a setting in which other papers report having estimation or identification problems. In this setting, suppose all $q_{j,t} \approx \eta$ (a small number), so $1 - q_{j,t} \approx 1$ and $1 - \frac{1}{2}q_{j,t} \approx 1$. Then, $s_t^{(n)} \approx 4L_t \frac{\sum_{j=1}^{n} d_{j,t}\eta}{\sum_{j=1}^{n} d_{j,t}} \approx 4L_t\eta$, and thus L_t and η are not point identified.

4 Robust Inference under Weak Identification

In this section, I describe the weak-identification robust econometric tools that are used in this paper, and I describe how these theoretical results can be used to construct confidence sets for LGD.

[22] Intuitively, it is also easy to understand why the model is not point-identified in this region. A non-flat CDS term structure contains different information about LGD and PD at each maturity point on the curve, so identification is generated by this information. When the term structure of both LGD and PD are flat, then the CDS term structure is flat, and additional maturity points beyond the shortest one do not add any additional information, so we cannot distinguish the respective contributions of LGD and PD to the CDS price using information from only one maturity point. However, set-identification comes from the fact that a certain level for the CDS spread $s_t^{(n)}$ will guarantee that PD and LGD cannot be too low. The identified sets of (L_t, q_t) are decreasing in size (in a nested sense) with the level of s_t.

4.1 S-test robust to weak identification

I treat observed CDS spreads as noisy realizations of the true price:

$$s_{it}^{(n)} = L_t \times g(\beta_{1t}, \beta_{2t}, \beta_{3t}; n) + \varepsilon_{it} \equiv h(L_t, \boldsymbol{\beta_t}; n) + \varepsilon_{it}, \text{ for all days } i \text{ in month } t. \tag{7}$$

If the model were globally identified, nonlinear least squares would be a straightforward choice for the estimation method. However, as described in the previous section, standard tests are not reliable in settings with weak identification, and robust estimation methods should be used instead.

Since the development of GMM in Hansen (1982), many papers have studied GMM under nonstandard conditions, with weak identification attracting considerable attention. Literature on inference for general nonlinear models under weak identification is not as extensive as the literature for robust linear instrumental variables estimation, but under the weak instrumental variables framework in nonlinear GMM, Stock and Wright (2000) derive the asymptotic distribution for the CUGMM objective function under very weak conditions, and estimate robust confidence intervals for the CRRA risk aversion coefficient in the consumption CAPM, which is weakly identified in the model Euler equations. Even though the weak identification in this model is not equivalent to that of the weak IV setting, the CUGMM results in Stock and Wright are sufficiently general that I can borrow the same methodology to construct confidence sets for LGD that are robust to weak identification.

To employ the methodology, the nonlinear least squares problem is recast in the GMM framework (as minimum distance estimation). From here onwards, I suppress the t subscripts for brevity, but it is understood that the model is estimated for each month t using the pooled daily term structure spreads from that month. Let $\theta = (L, \boldsymbol{\beta}) = (L, \beta_1, \beta_2, \beta_3)$ be the parameters of the model, and take the score function of the least squares problem as the estimating equations:

$$\phi_i(\theta) = \phi_i(\theta; s) = \left(s_i^{(n)} - h(\theta; n)\right) \frac{\partial h(\theta; n)}{\partial \theta}. \tag{8}$$

Further, I define notation for the standardized moment $\Psi_N(\theta) = N^{-\frac{1}{2}} \sum_{i=1}^{N} [\phi_i(\theta) - E\phi_i(\theta)]$ and the asymptotic variance of the standardized moment, $\Omega(\theta) = lim_{N\to\infty} E\Psi_N(\theta)\Psi_N(\theta)'$.

Then, the standard CUGMM objective function is

$$S_{cN}(\theta) = \left[N^{-\frac{1}{2}} \sum_{i=1}^{N} \phi_i(\theta)\right]' W_N(\theta) \left[N^{-\frac{1}{2}} \sum_{i=1}^{N} \phi_i(\theta)\right], \tag{9}$$

where $W_N(\theta)$ is an $O_p(1)$ positive definite 4×4 weighting matrix that is a function of θ.

Suppose $\theta_0 = (L_0, \boldsymbol{\beta}_0)$ are the true parameters of the model. The result I use from Stock and Wright only requires the weak assumptions that the GMM moment condition obeys the central limit theorem locally at θ_0, i.e. $\Psi_N(\theta_0) \xrightarrow{d} N(0, \Omega(\theta_0))$,[23] and that the weighting matrix used is consistent for the inverse of the asymptotic variance of the standardized moment at θ_0, i.e. $W_N(\theta_0) \xrightarrow{p} \Omega(\theta_0)^{-1}$. Under these assumptions, Theorem 2 in Stock and Wright shows that $S_{cN}(\theta_0) \xrightarrow{d} \chi^2_4$.

From this result, one can obtain an asymptotic level-α confidence set for θ, called S-sets in Stock and Wright (2000), by inverting the CUGMM objection function surface, that is

$$\mathcal{C}_{\theta,\alpha} = \left\{ \theta | S_{cN}(\theta) \leq \chi^2_4(1-\alpha) \right\}, \tag{10}$$

where $\chi^2_4(1-\alpha)$ is the $100(1-\alpha)\%$ critical value of the χ^2_4 distribution. However, constructing $\mathcal{C}_\theta$ empirically involves extensive computation; specifically, $S_{cN}(\theta)$ must be evaluated over a very fine grid of values of θ spanning the 4-dimensional parameter space. Instead, given our main focus is estimating LGD, we use a result similar to Theorem 3 in Stock and Wright, which allows for confidence sets of a subset of the parameters.[24]

Under the two assumptions on the GMM moments and weighting matrix, we get that the profile CUGMM objective function evaluated at the true LGD value L_0 converges in distribution to a chi-square distribution, i.e. $S_cT(L) \equiv S_{cT}(L_0, \hat{\boldsymbol{\beta}}(L_0)) \xrightarrow{d} \chi^2_1$, where $\hat{\boldsymbol{\beta}}(\tilde{L}) = argmin_{\boldsymbol{\beta}} S_{c_T}(\tilde{L}, \boldsymbol{\beta})$. This result allows us to obtain asymptotic level-α confidence sets for L by inverting the profile S-function, i.e., the set $\mathcal{C}_{L,\alpha} = \left\{ L | S_{cT}(L, \hat{\boldsymbol{\beta}}(L)) \leq \chi^2_1(1-\alpha) \right\}$. Equivalently, this procedure can be thought of a test (called an S-test) of model specification and parameter identification: the model with parameter value $\tilde{L} \in (0,1]$ is rejected at the $(1-\alpha)$ confidence level if the S-function evaluated at $\tilde{L}$ is greater than the chi-square critical value. If all values in $(0,1]$ are rejected, then the model has been rejected entirely.

Note that while the computation time is drastically reduced compared to estimating confidence sets for θ, the procedure, which will be described in further detail in the next subsection, still requires mapping out the profile-S curve for many values of L, and each

[23] Among other scenarios (structural breaks, existence of higher moments, etc), this assumption precludes the sequence of moment conditions from being integrated of order one or higher, which is satisfied in this application.

[24] The relationship between the strongly identified (β_2, β_3) and weakly identified (L, β_1) parameters in this CDS model is not characterized in the exact manner of Assumption C of Stock and Wright (2000). Also, note that in the weak IV model studied in Stock and Wright (2000), Theorem 3 does not allow for asymptotically valid confidence sets for *a subset of* the weakly identified parameters. However, Theorem 3 holds since when $L = L_0$, $(\beta_1, \beta_2, \beta_3)$ are well-identified (and thus consistently estimated), which is the general purpose of requiring Assumption C.

point on this curve is obtained from a nonlinear optimization over $\boldsymbol{\beta}$.

The asymptotic result from which the S-sets are developed relies jointly on identification of the true parameter θ_0 and validity of the GMM orthogonality conditions. Therefore, as alluded to above, the confidence set consists of parameter values at which the joint hypothesis that $\theta = \theta_0$ *and* $E[\phi_t(\theta; s)] = 0$ is unable to be rejected. In contrast, conventional tests (Wald, Likelihood Ratio) operate under the assumption that the orthogonality conditions hold, and are only a test of parameter identification.[25]

This feature of the S-sets is valuable as it can detect a misspecified model by rejecting the entire parameter space. However, the trade off is that non-empty confidence sets require some care in interpretation that is not necessary with conventional tests. If an S-set is nonempty, there are two possibilities: either the model is correctly specified and parameter estimates are given by the S-set, or the model is misspecified, but, for the values in the S-set, the test lacks power to reject the model. Very large S-sets suggest that there is little evidence to distinguish between parameter values, but small confidence sets are more difficult to interpret; they could reflect a correctly specified and precisely estimated model, or a misspecified model that the test was not powerful enough to reject for a small set of parameter values. There is no solution to this problem, which is faced by all tests of this type, but in the results section, in addition to running the 0.05-level test, I also consider the more powerful 0.10-level test and find that results do not change much.[26]

4.2 Estimation procedure for confidence sets for LGD

As described in Section 3, I assume that CDS spreads from each calendar month are generated from the same data-generating process (DGP), so the model can be estimated independently each month using the available panel of observed CDS spreads. Generally, this is a reasonable assumption since CDS spread dynamics are very slow-moving as shown in Section 2.2, but it is possible that for some months, the constant parameter assumption is not a good approximation. For example, we could envision a news shock in the middle of some month that causes sudden changes in beliefs about a firm's credit. However, the important point is that the S-sets are valid regardless of whether the assumption holds because they jointly test parameter identification *and* model specification. Thus, if the constant model assumption were false, the model would be rejected for all values of

[25] As discussed in Stock and Wright, under conventional asymptotics (not in the presence of weak-identification), $S_{cT}(\theta_0)$ is asymptotically the sum of the Likelihood Ratio (LR) statistic testing $\theta = \theta_0$ and Hansen's (1982) J statistic testing the over-identifying conditions.

[26] See Table A.2 for 0.10-level results.

LGD, i.e. S-sets would be empty.[27]

There are 4214 firm-months in total, and for each firm-month, the model is estimated using approximately 176 (approximately 22 days$\times$8 contracts/day) pooled individual CDS spreads. The weighting matrix used is the inverse of an asymptotic variance estimator that accounts for maturity point effects by clustering and serial correlation through a Newey-West HAC style kernel.

For each month, a confidence set for LGD can be obtained by evaluating the profile S-function $S_{cT}(L, \hat{\boldsymbol{\beta}}(L))$ for a grid of values $\mathcal{L}$ for LGD and comparing the resulting values with a chi-square critical value. If the profile S-function evaluated at L is less than the critical value, then L is in the confidence set. Note that the confidence set may be disjoint. As an initial grid for LGD, I use $\mathcal{L} = \{0.01, 0.02, ...0.98, 0.99, 1\}$. For each value of L, $S_{cT}(L, \boldsymbol{\beta})$ is minimized over $\boldsymbol{\beta}$, with parameters constrained so that default probability curves are in [0,1].[28]

In the second stage, I evaluate $S_{cT}(L, \hat{\boldsymbol{\beta}}(L))$ on a finer grid (grid space of 0.001 instead of 0.01 as in the original grid $\mathcal{L}$) around the values of L where $S_{cT}(L, \hat{\boldsymbol{\beta}}(L))$ attains a local minima or is close to the chi-square critical value. Respectively, adding these two finer grids reduce the chance of incorrectly rejecting the model (finding an empty S-set) due to errors from LGD grid discreteness and allows estimation of the S-set end-points to be precise to 0.001.

4.3 Examples of Profile S-functions and S-sets

Before presenting the main results of the paper, I plot, in Figure 4, concentrated S-functions from 4 months to illustrate S-sets constructed from actual data. In addition, I describe how the confidence sets from this method can look different from conventional confidence intervals from likelihood ratio (LR) and Wald tests. A confidence set from a standard Wald

[27] As a robustness check, I also assume a constant CDS model over longer and shorter time periods (semi-monthly and quarterly) and check for changes in estimated confidence sets. The detailed results are presented in the appendix, but generally there are two contrasting effects from changing the data aggregation period: using less data (using CDS data over a shorter period) will reduce the power of the test, but on the other hand, a longer time period is more likely to compromise the assumption of the constant DGP, leading to either larger confidence sets due to "extra noise" in the data or empty confidence sets if there is enough power to outright reject the model.

[28] Due to the nonlinearity of the profile S-function, I start the optimization procedure from around 1000 initial values from the compact parameter space for $\boldsymbol{\beta}$. These values were not sampled uniformly from the entire parameter space. Instead, more "reasonable" regions (those mapping to lower default probabilities) were sampled more frequently. In addition, if the profile S-curve (as a function of L) has upward spikes in it, indicating that the first round of optimization stopped at a local minima, the optimization procedure for those values of L was repeated using initial values similar to the $\hat{\boldsymbol{\beta}}$ estimates from nearby values of L. This was repeated until the spikes in the profile-objective curve were resolved.

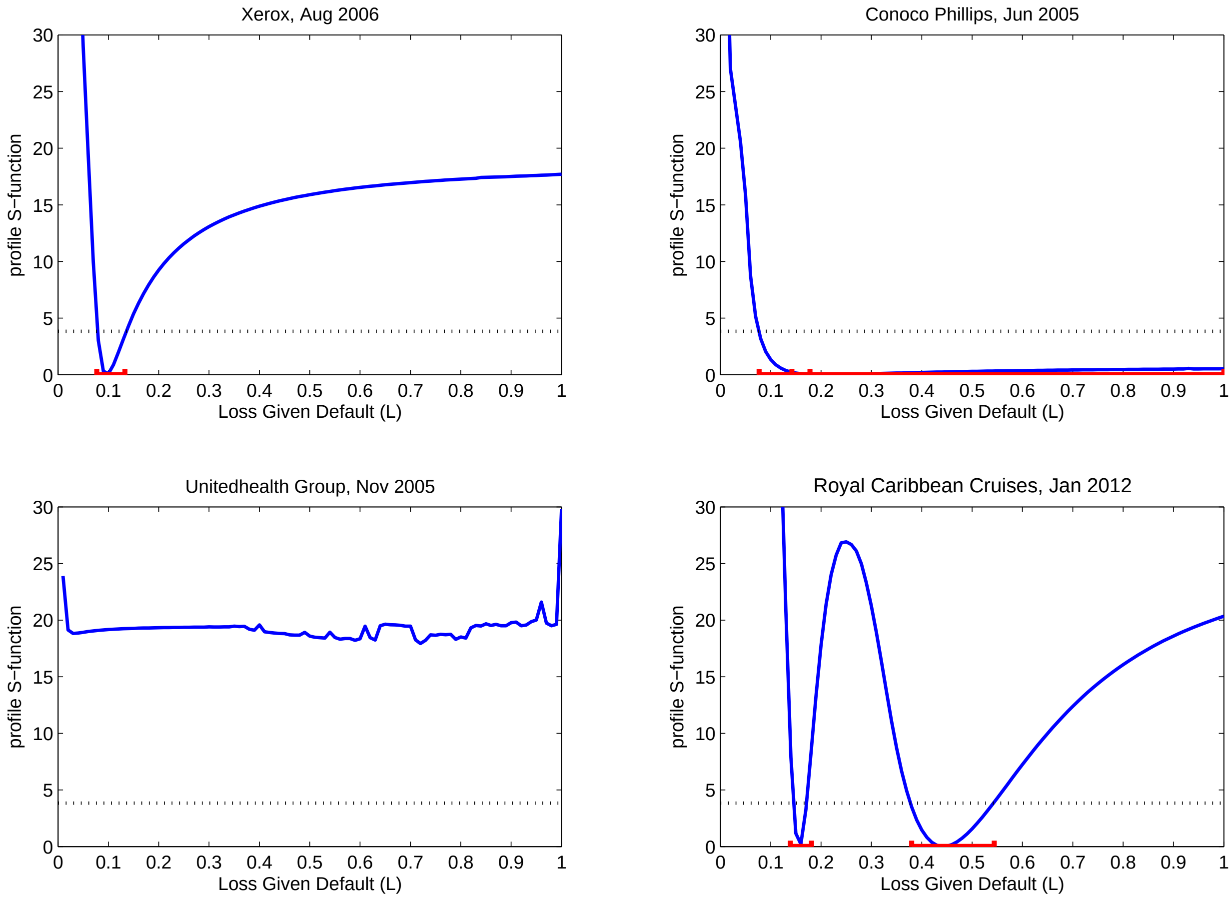


Figure 4: *This figure presents profile S-functions for four representative firm-months and depicts the estimated confidence sets for loss given default in red brackets on the x-axis. The black dotted line represents the 95% chi-square critical value cutoff used to construct the confidence sets.*

test is generally a symmetric interval around the point estimate, and confidence sets from both LR and Wald tests are not empty and not disjoint by construction. The two top subplots show typical shapes of the profile S-curve that yield small and very large confidence sets. The confidence sets are not symmetric around the CUGMM point estimate (CUE), where the profile S-function reaches its minima. In addition, the bottom left subplot shows an firm-month for which an empty confidence set is estimated. The bottom right subplot shows a disjoint confidence set. Disjoint S-sets are observed for 349 out of 4214 firm-months in this sample, and most of them cover a large portion of the (0,1] parameter space. Finally, we observe that the profile S-function diverges for very small values of LGD, which is expected since the CDS price is undefined when LGD equals zero.

5 Empirical Results from Joint Estimation

In this section, I present the main results. In particular, I describe the estimated confidence sets for LGD, and address the two main questions of the paper: can LGD be precisely estimated when jointly estimated with PD in this term structure model, and if so, is 60% a good estimate for LGD? In addition, I explore the differences in the results across reference entities and over the sample period, and investigate whether characteristics of the monthly CDS data affect how precisely LGD is estimated. Finally, I study the implied default curves that are jointly estimated with LGD, and discuss the implications on credit risk premia.

5.1 Confidence Set Lengths and Locations

Here, I present a summary of the sizes and location (on the (0,1] parameter space) of the 4214 estimated monthly confidence sets for LGD and address a main question of whether LGD can be jointly estimated alongside default probabilities, which in effect is asking how large are the confidence sets for LGD? If the estimated confidence sets are all large subsets of the total parameter space (0,1], then it implies that many combinations of LGD and PD are indistinguishable in their ability to fit the spread data, and hence, we gain little information about the true value of risk-neutral LGD.

In Table 2, I divide the 4214 S-sets into bins based on their size, and report the proportion of confidence sets in each bin. I find that the confidence sets for LGD are concentrated in the smallest and largest bins. For 46.7% of firm-months, LGD is estimated very precisely, with confidence sets less than 0.10 in length. Another 7.0% of firm-months have confidence sets between 0.10 and 0.20 in length. However, a large fraction (28.7%) of firm-months have estimated confidence sets that are very large, greater than 0.80 in length, which indicates an almost complete lack of information in the data to distinguish between the roles of LGD and PD in the cross-sectional model. Also, it is interesting to note that there are relatively few firm-months in the mid-sized confidence set bins – only 13.5% of confidence sets are between 0.20 and 0.80 in length.

Table 2: Summary of the length and location of LGD confidence sets (all firms)

	Proportion (%)	avg C_L center
Confidence set length:		
(0,0.1]	46.7	0.12
(0.1,0.2]	7.0	0.20
(0.2,0.4]	4.4	0.43
(0.4,0.6]	4.2	0.58
(0.6,0.8]	4.9	0.56
(0.8,1]	28.7	0.53
empty	4.0	–
Testing $L_0 = 0.60$:		
$0.60 \in C_L$	37.1	
$0.60 \in C_L \| \ell(C_L) < 0.5$	2.9	
$0.60 \in C_L \| \ell(C_L) < 0.2$	≤ 0.1	

Notes: This table groups the 4214 monthly LGD confidence sets into bins by their length, and presents the proportion of confidence sets belonging to each bin. The bottom panel presents the proportion of confidence sets that include 0.60, out of all 4214 firm-months and for the subsets of firm-months with confidence sets less than 0.5 and 0.2 in length.

In addition, 4.0% of the confidence sets are empty, meaning that the model is rejected for all values of LGD at the 0.05 level. Rejection of the model could be due Type 1 error, or to one or more of the following: the Nelson-Siegel curve parameterization for PD is incorrect, the flat term-structure for LGD is a poor fit, or the constant parameter assumption failed for the month in question.[29]

The right-most column of Table 2 characterizes the bins of S-sets along an additional dimension — where they are located in the (0,1] parameter space. By construction, the average midpoint of the largest confidence sets must be around 0.50. For shorter intervals, the midpoint can lie (almost) anywhere on (0,1]. However, I find that when LGD is precisely estimated, the estimated values are very low. The smallest group of S-sets is centered at 0.12 on average, and the second smallest group (length between 0.10–0.20) is centered at 0.20.

[29] I also construct S-sets for LGD with type one error of 0.10 for a more powerful test, and find very little difference in the results (see appendix table A.2 for a summary of the distribution of S-set sizes). The proportion of small S-sets increases 5% and the proportion of empty sets increases 1%.

5.2 Is loss given default really 0.60?

The second main question I address is whether 0.60 is a good approximation for LGD. From the results in the previous subsection, we already infer that 0.60 is not suitable for the many firm-months in which low values of LGD are estimated. In fact, the estimated S-sets are a formal test for this point specification test; the null hypothesis $L_0 = 0.60$ is rejected (at the 0.05 level) for a given month if and only if 0.60 is not in the monthly 95% S-set.

In the bottom panel of Table 2, I present the proportion of confidence sets that include 0.60 for the entire set of firm-months, and also for sets of months where LGD is more precisely estimated. We observe that over all firms and all months, 37.1% of confidence sets do not include 0.60. More strikingly, if only considering months with confidence sets that are less than 0.5 in length, only 2.9% of these confidence sets contain 0.60, and less than 0.1% of confidence sets less than 0.2 in length include the value 0.60. Thus we conclude that in the months that LGD is precisely estimated, 0.60 is rejected as a value for risk neutral LGD. However, these results do not rule out that 0.60 is an appropriate value for LGD in the more than one-third of firm-months that have very large LGD confidence sets.

5.3 Confidence sets across reference entities and over time

The above analysis focuses on characterizing the aggregate collection of estimated confidence sets, but we are also interested in discovering heterogeneity in the results across firms and over time. The top two subplots of Figure 5 summarize the distribution of confidence set sizes for each of the 30 reference entities. The top panel plots the proportion of S-sets that are less than 0.1 in length and the average S-set center of these small sets, per firm. The middle panel plots the proportion of large S-sets (length greater than 0.6), and the proportion of empty confidence sets (to make the plot easier to read, the firms have been ordered by the proportion of small confidence sets). We see that the proportion of small confidence sets varies substantially across firms from 66% (Arrow Electronics, ARW) to only 21% (Union Pacific Corp., UNP). We also see some variation in the average midpoints of these small S-sets; average midpoints lie between 0.05 (UNP) and 0.21 (Goodyear Tire, GT). These results lead us to ask whether there are variables that can explain the differences in confidence set size and in the estimated level of LGD when precisely estimated. We investigate the former in section 5.4 and the latter in section 5.7 after describing the estimates for the forward default probability curve.

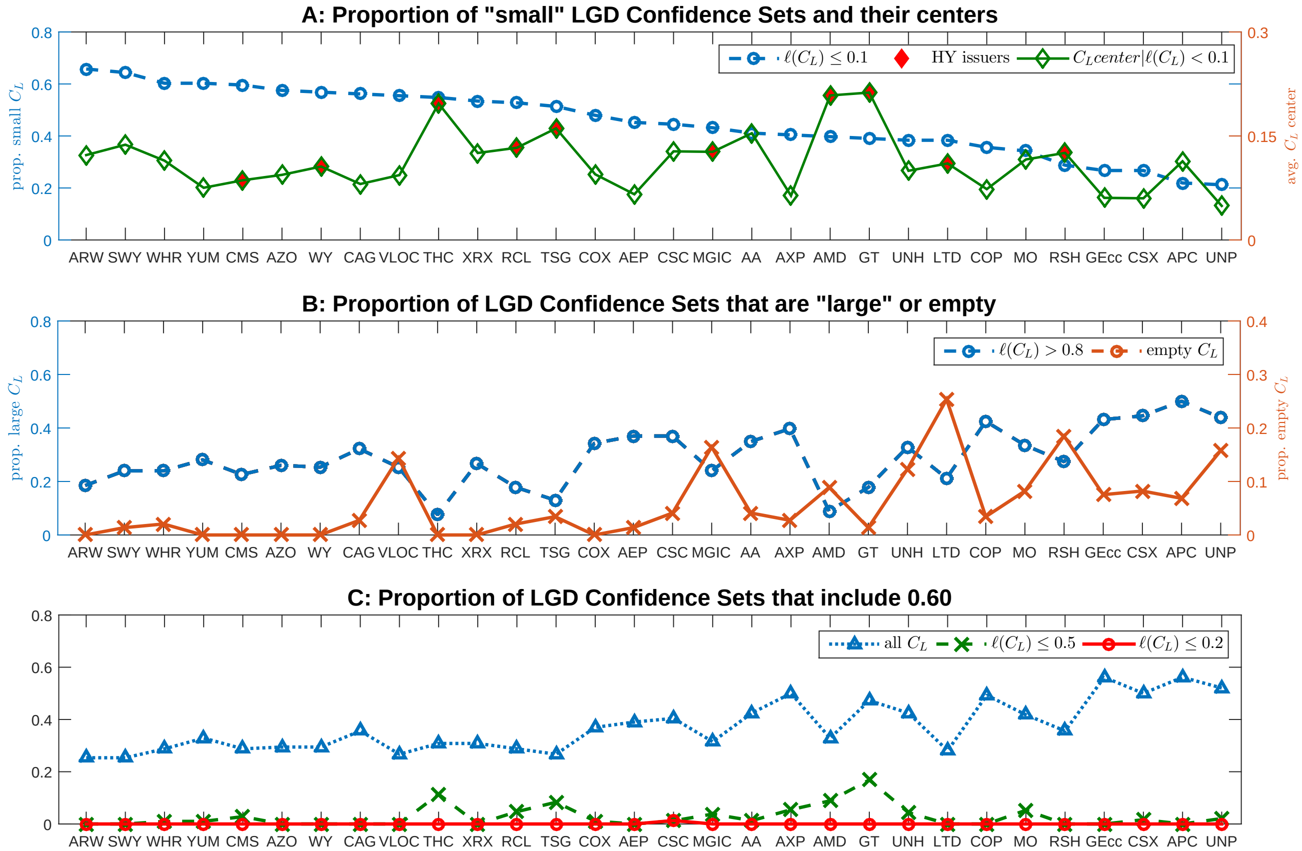


Figure 5: *This figure presents summaries of the monthly confidence sets for each of the 30 reference entities. The top panel of this figure plots (1) the proportion of months (out of 146 months for most assets) that have estimated confidence sets with length less than 0.1 with circle markers, and (2) the midpoints (centers) of the small confidence sets with diamond markers. The middle panel plots (1) the proportion of "large" confidence sets (length greater than 0.8) with circle markers and (2) the proportion of empty sets with x's. The bottom plot presents the proportion of LGD confidence sets that include the value of 0.60, for all confidence sets and among the subset that have length less than 0.5 and 0.2.*

Additionally, we can investigate if 0.60 is a better estimate for LGD for certain reference entities. The bottom panel of Figure 5 plots the proportion of confidence sets that include 0.60, keeping the same order of firms as in the top panel. This proportion ranges from 25% for Arrow Electronics (ARW) and Safeway (SWY) to 56% for GE-Capital Corp (GEc) and Anadarko Petroleum Corp. (APC). In addition, the bottom panel also presents the proportions of confidence sets with length less than 0.5 or less than 0.2 that include the value 0.60. The latter set of proportions ranges from 0 to 0.1%, which further illustrates that 0.60 is generally only in large S-sets, i.e., when there is insufficient information in the data to jointly estimate LGD and PD.

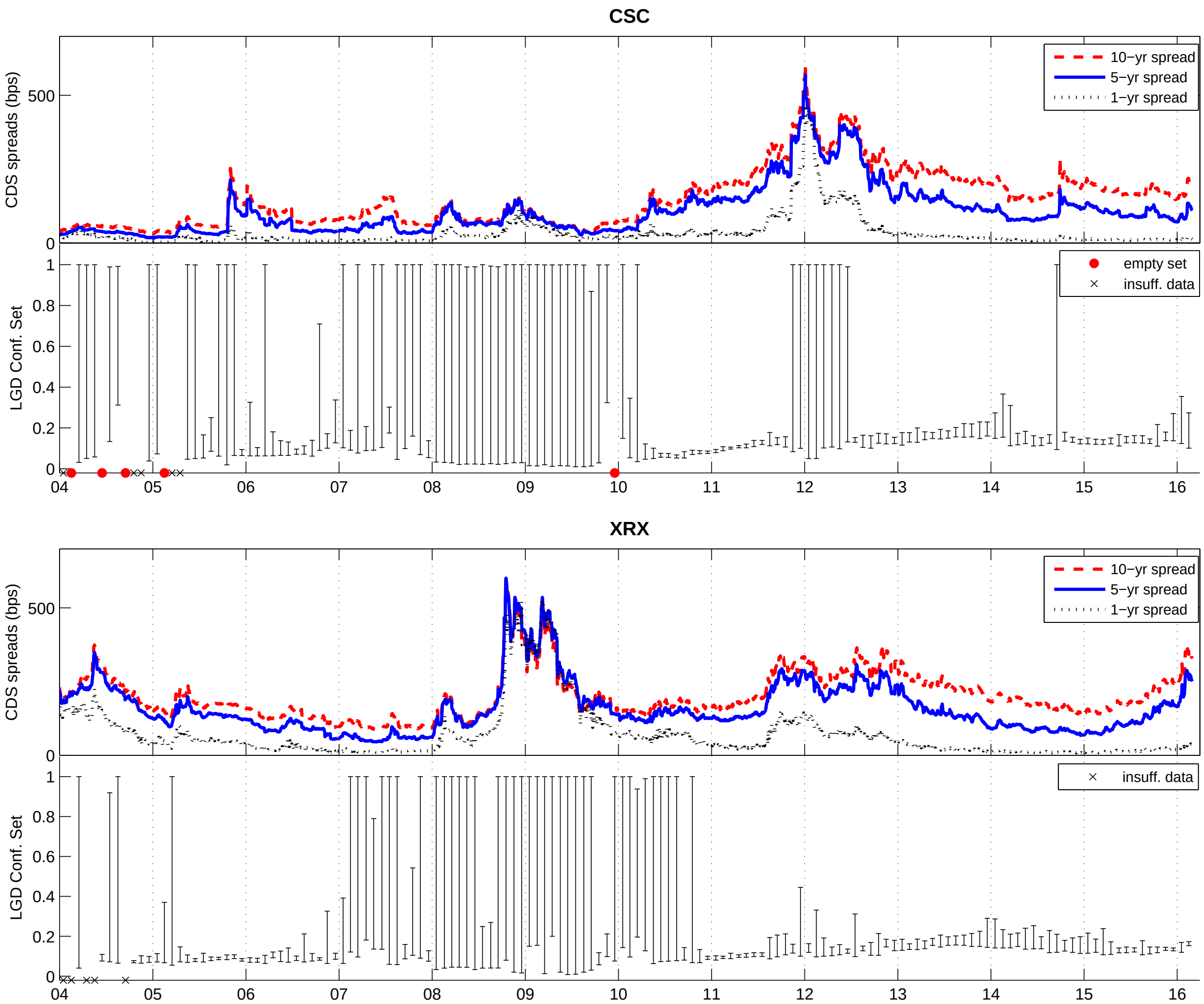


Figure 6: *This figure presents all estimated loss given default confidence sets (S-sets) for two reference entities, Computer Sciences Corporation (top) and Xerox Corporation (bottom). Empty confidence sets are denoted with a red filled dot; months that were dropped due to insufficient data are marked with 'x'. In addition, the 1, 5, and 10-year daily CDS spreads are plotted above the confidence sets for comparison and context.*

In addition to tabulating summaries of the confidence sets for individual assets, we can plot the S-sets over time, per firm. Figure 6 plots the confidence sets for two representative firms, Computer Sciences Corp. (CSC) and Xerox (XRX). Both plots show a pattern that is present for most all issuers: almost all of the confidence sets from the financial crisis period (late 2007 through 2009) are large, while the periods before and after have a smaller proportion of large confidence sets. Another common observation is that firms may have specific time periods outside of the financial crisis period during which CDS spreads peak and estimated LGD confidence sets are large, e.g., November 2011–June 2012 for CSC.

To explore this further, I divide the 146 month sample period into four sub-periods of approximately equal length, and in Table 3, I reproduce the confidence set summaries as in Table 2 for each sub-period. Period 1 runs from Jan 2004–Dec 2006 (36 months), period 2 runs from Jan 2007–Dec 2009 (36 months), period 3 runs from Jan 2010–Dec 2012 (36 months), and period 4 runs from Jan 2013–Feb 2016 (38 months). Note that periods 1 and 2 are mostly in the pre CDS Big Bang era, while periods 3 and 4 are post-CDS Big Bang.[30] Across these four sub-periods, the estimated LGD S-sets are quite different. In the first and third sub-periods, around 40–55% of the firm-months have confidence sets with length less than 0.10, and 20–30% of confidence sets are larger than 0.8 in length. In contrast, during the second sub-period, which includes the Global Financial Crisis, only 17.2% of confidence sets had length less than 0.10 and 56.0% of confidence sets had length greater than 0.8, while in the most recent sub-period, 72.1% of confidence sets were "small", and only 8.8% were "large". In terms of empty confidence sets, the third sub-period has the fewest at only 1.7%, while the second sub-period has the largest proportion at 6.0%, however, all values are relatively close to the specified level of the test. So the main pattern observed is that the confidence sets are most "informative" in the recent subperiod, and least "informative" during the Global Financial Crisis.

[30] The CDS Big Bang was announced in April 2009, and policies were implemented 2-3 months later.

Table 3: Summary of the length and location of Confidence Sets, by subperiod (all firms)

	Jan04–Dec06		Jan07–Dec09		Jan10–Dec12		Jan13–Feb16	
	prop. (%)	avg C_L center	prop. (%)	avg C_L center	prop. (%)	avg C_L center	prop. (%)	avg C_L center
Confidence set length:								
≤ 0.1	42.7	0.08	17.2	0.11	54.3	0.11	72.1	0.14
(0.1,0.2]	9.1	0.14	6.0	0.22	6.3	0.25	6.9	0.19
(0.2,0.4]	4.4	0.34	4.8	0.38	5.8	0.56	2.7	0.41
(0.4,0.6]	3.9	0.51	3.7	0.58	7.5	0.63	1.8	0.52
(0.6,0.8]	6.0	0.52	6.2	0.57	3.9	0.59	3.5	0.58
(0.8,1]	29.9	0.52	56.0	0.52	20.6	0.53	8.8	0.55
empty	3.9	–	6.0	–	1.7	–	4.2	–
Testing $L_0 = 0.60$:								
$0.60 \in C_L$	39.1	–	65.4	–	31.1	–	13.1	–
$0.60 \in C_L \| \ell(C_L) < 0.5$	2.2	–	5.3	–	5.1	–	0.7	–
$0.60 \in C_L \| \ell(C_L) < 0.2$	0.0	–	0.4	–	0.0	–	0.0	–

Notes: This top portion of this table groups the monthly confidence sets in bins by their length, and presents the proportion belonging to each bin and the average confidence set center per bin. The bottom three rows summarize the proportion of confidence sets that include 0.60 for all firm-months and the subsets of firm-months with confidence set length less than 0.5 or 0.2.

A final comment on the confidence sets over time: it is observed that the small confidence intervals are usually estimated clustered together; however, there are numerous cases when a string of small monthly confidence intervals, roughly estimated at the same level, is interrupted by one or several large confidence intervals that span almost the entire parameter space (for example, the second half of 2014 for Computer Sciences Corp. in Figure 6). I stress that this does not necessarily mean that risk neutral LGD jumped from around 0.10 to 0.90 (or even to the midpoint of the large confidence set) and then back, over the span of one month. This sequence of confidence intervals is also consistent with the true implied LGD staying steady at 0.10 over all the months, as 0.10 is in the confidence set for all three months. The large S-set is just a consequence of the data containing limited information that month, so virtually no values in the LGD parameter space could be rejected. In this next subsection, I will explore whether characteristics of the CDS spread data can explain differences in sizes of the confidence sets.

5.4 Effects of CDS data characteristics on Confidence Set Size

In this section, I summarize how three characteristics of monthly CDS data are related to the size of LGD confidence sets, using a fractional-response general linear regression (with logit link function) by regressing the size of monthly confidence sets on explanatory variables constructed from monthly CDS data.

In section 3.4, it is shown that when the PD term structure is flat, the cross sectional model is not identified. Maintaining the model assumption that the LGD term structure is flat, flat PD term structure implies a theoretical CDS term structure that is also flat. Thus, I investigate whether LGD S-sets are larger for months with flatter CDS term structures by including in the regression a measure of CDS term structure flatness, or rather "unflatness", as the measure is the average absolute deviation of the 7 CDS spreads of maturities 6-month, 1–4, 7 and 10 years, from the 5-year CDS spread. In addition, noisy data caused by large errors diminishes the test's ability to extract information about the true model, so I also include a measure of noise variance – the within month coefficient of variation of 5-year spreads. The coefficient of variation is defined as the sample standard deviation divided by the sample mean, so more precisely it is a noise-to-signal ratio. Finally, I also include the monthly sample mean of the 5-year spreads as an explanatory variable, which can be viewed as a proxy for level of credit distress of the reference entity, and is also related to identification, because as noted in section 3.4, if spreads are high when PD term structure is flat, then the identified set for $(L_t, q_t) = (L_t, \beta_{1,t})$ is smaller than when spreads are low. Also, point identification may be threatened when CDS spreads are near zero.[31,32]

The regression results are reported in Table 4.[33] The regression is conducted over the full sample and separately for each subperiod. For each sample period, explanatory variables have a strong relationship with the confidence set length, and have signs in the expected direction. When the CDS term structure is less flat (i.e., the "unflatness" variable

[31] An alternate regression specification using the standard deviation of 5-year spreads (unstandardized) as the "noise" variable produces similar results, however the three explanatory variables are highly collinear, whereas the coefficient of variation displays very low correlation with both average 5-year CDS spread and term-structure flatness (though, average 5-year spread and the unflatness measure have a sample correlation of 0.79 over all firms and all months). In addition, the coefficient of variation measure is more comparable in level across all firms, whereas the average standard deviation across firms is substantially varied.

[32] The estimates of β_2 and β_3 could be used to used as explanatory variables, however I choose to only use variables that can be directly measured from the CDS spreads before any estimation takes place.

[33] I stress that this regression is merely a summary tool for exploring how characteristics of monthly CDS spreads affect S-set size. It is not proposed as a model for the dependent variable (S-set length), nor is this variable truly intended to be interpreted as a random variable. It follows that the standard errors are not interpretable in the regular sense, but they are reported as they still do provide some indication of how strong the regression association is.

takes a higher value), LGD is more precisely estimated (smaller S-sets). Also, a higher coefficient of variation, a proxy for noisier data, is associated with a larger confidence set, which is also expected since information in the data is more heavily contaminated by noise. Finally, the level of the 5-year CDS has a positive effect on confidence set length. In the full sample regression, it is possible that the significance of the coefficient estimates are driven by the large differences in CDS spread behavior during the "financial crisis" period and in the rest of the months. However, the results are robust for all four subperiods, as shown in Table 4, and Periods 1, 3, and 4 do not overlap with the financial crisis, so the significance of these three credit spread factors is not only driven by the disparate financial crisis period.

Table 4: Summary regression on confidence set length

	Full sample	Jan04–Dec06	Jan07–Dec09	Jan10–Dec12	Jan13–Feb16
constant	**-1.21**	**-0.67**	-0.11	**-2.02**	**-1.97**
	(0.08)	**(0.16)**	(0.16)	**(0.18)**	**(0.17)**
avg(CDS5yr)	**22.70**	**132.25**	-0.92	**23.74**	**27.72**
	(2.60)	**(20.64)**	(3.52)	**(6.95)**	**(5.89)**
coef.var.(CDS5yr)	**18.97**	**19.83**	**14.18**	**24.00**	**14.06**
	(0.98)	**(2.44)**	**(1.67)**	**(2.36)**	**(2.04)**
"unflatness"	**-180.07**	**-668.78**	**-100.52**	**-108.95**	**-145.77**
	(13.79)	**(86.96)**	**(26.35)**	**(30.64)**	**(28.54)**
R-square	0.32	0.29	0.23	0.26	0.22

Notes: This table presents results from fractional response logit regressions of confidence set length on explanatory variables constructed from monthly CDS data. Coefficient estimates that are significant at the 0.05-level are displayed in bold font. These regressions are limited to firm-months where the LGD is not empty.

5.5 Estimated default probability term structure

As explained in Section 4.1, I do not construct confidence sets for the full parameter vector $(L, \beta_1, \beta_2, \beta_3)$ in this paper due to computational burden, however, we can still study the (point) estimated PD term structures given a fixed value of L, i.e. the PD term structure $\{q_j(\hat{\beta}(L))\}_{j=1,\ldots n}$, where $\hat{\beta}(L) = argmin_\beta S_{cT}(L, \beta)$.[34] In particular, I am interested in comparing the estimated PD term structure corresponding to LGD fixed at

[34] Note that the values of LGD in the S-sets and their corresponding estimated PD term structures, i.e. $\{(L, \hat{\beta}(L)) : L \in C_L\}$, represent estimated models that are not rejected under the S-test with level 0.05, however the set of β estimates $\{\hat{\beta}(L) : L \in C_L\}$ do not make up a correctly sized (asymptotically) confidence set for $(\beta_1, \beta_2, \beta_3)$.

0.60, $\{q_j(\hat{\beta}(0.60))\}_j$, to the PD term structure when L is the CUGMM point estimate, i.e. $\{q_j(\hat{\beta}(L_{cue}))\}_j$, where L_{cue} is the CUGMM point estimate for L. For brevity, I will refer to $\{q_j(\hat{\beta}(0.60))\}_j$ as the "conventional PD" and $\{q_j(\hat{\beta}(L_{cue}))\}_j$ as the "estimated PD".

I limit this analysis to months when LGD is precisely estimated (S-set length is less than 0.1), which is the case for 1970 firm-months out of the total 4214. I find that for the months with small confidence sets, the estimated PD term structure can be approximately 4-20 times higher than the conventional PD term structure, and is, on average, 8.6 times higher than conventional PD. Table 5 further distinguishes between investment grade and high yield reference entities, and presents average values of estimated and conventional forward default probabilities and the average ratio of estimated to conventional PD for four maturity points. The average point estimates of LGD are 0.092 for investment grade issuers and 0.141 for high yield issuers, which are, on average, 6.5 and 4.3 times lower than 0.60, respectively. On average, the term structure of estimated PD is around 9-12 times higher than that of conventional PD for investment grade entities, and 5-7 times higher for high yield entities.

I also include, in Table 5, the average *cumulative* default probability at 4 points along the term structure, denoted by $Q_{j,t}$, and I note that the ratio between the estimated cumulative PD and conventional cumulative PD decreases as the maturity increases, i.e., as default probabilities are compounded. However, even at 10 years, the estimated cumulative default rate is 4.3 times higher for investment grade issuers and 2.3 times higher for high yield issuers, a substantial magnitude of difference. Therefore, for the months where LGD is precisely estimated, default probabilities are hugely biased when LGD is fixed at 0.60.

Table 5: Risk-neutral default probabilities when LGD is estimated or fixed exogenously at 0.60

	$q_{1,t}$	$q_{4,t}$	$q_{20,t}$	$q_{40,t}$	$Q_{1,t}$	$Q_{4,t}$	$Q_{20,t}$	$Q_{40,t}$	avg λ_q
					Investment Grade				
avg PD(cue)	0.007	0.010	0.062	0.107	0.007	0.032	0.438	0.858	0.196
avg PD(0.6)	0.001	0.001	0.007	0.012	0.001	0.004	0.072	0.230	0.026
avg PD(cue)/PD(0.6)	11.6	8.8	10.0	10.6	11.6	9.3	7.0	4.3	7.5
					High Yield				
avg PD(cue)	0.018	0.024	0.100	0.163	0.018	0.055	0.589	0.943	0.286
avg PD(0.6)	0.004	0.006	0.019	0.028	0.004	0.017	0.185	0.461	0.062
avg PD(cue)/PD(0.6)	6.4	4.8	6.1	6.7	6.4	4.2	3.6	2.3	4.6

Notes: This table compares the average PD term structure (forward quarterly probabilities of default) and average cumulative PD term structure ($Q_{j,t}$ is the month-t expected (risk-neutral) cumulative probability of default through the j-th quarter from month t) corresponding to LGD fixed at 0.60 to the PD term structure at the CUGMM point estimate (labelled 'cue') for investment grade entities and high yield entities. The average of the ratios of these two term structures is also calculated. Analysis is limited to months when LGD is precisely estimated (when confidence set length is less than 0.1).

5.6 Discussion of risk premia implications

Existing papers that quantify risk premium in credit spreads use models with loss given default fixed exogenously, so in addition to possibly estimating distorted values, the entire premium (excess returns) in CDS is counted as compensation for default probability risk. For example, Driessen (2005) and Berndt et al. (2008) use the ratio of risk-neutral to actual default probabilities as a measure of the proportional premium for bearing default risk.[35] They fix LGD at 0.60 and approximately 0.75, respectively, and estimate the risk premium ratio of default intensities to be around 2–4 for U.S. corporates. However, there is uncertainty regarding loss given default as well, and an important question that arises is whether risk in LGD is priced, or if the premium observed in credit spreads is correctly attributed to the probability of default.

To investigate how the magnitude of risk premium on default probabilities might change when LGD is estimated from the data, I compute a simple risk-neutral default intensity loosely comparable to the default intensities estimated in Driessen (2005) and Berndt et al. (2008). I back out the risk-neutral default intensity associated with the cu-

[35] Berndt et al. (2008) gives the example that "if this ratio is 2.0 for a particular firm, date or maturity, then market-based insurance of default would be priced at roughly twice the expected discounted default loss." Driessen (2005) studies credit risk using corporate bond data.

mulative 10-year risk-neutral probability of default presented in Table 5, assuming constant intensity over the 10-year period. Unlike Driessen (2005) and Berndt et al. (2008), this paper does not contain the machinery to extract actual (objective) expectations of default intensity, so I cannot compute a value for the proportional risk premium measure. However, given the convenient form of the proportional premium measure as a ratio of risk-neutral to objective default intensities, we know that since the default intensity under joint LGD-PD estimation is around 7.5 times higher than the default intensity under LGD fixed at 0.60 for IG firms, then the proportional risk premium under joint estimation is thus 7.5 times higher than the proportional risk premium under the assumption that LGD = 0.60. I also calculate that the average ratio of the default intensities is 4.6 for HY firms, meaning that the proportional risk premium would be 4.6 timees higher under joint estimation than if LGD were fixed exogenously at 0.60. So in summary, the risk premium on default probability is grossly understated when LGD is fixed at 0.60 rather than estimated.

A puzzle that now arises is that the risk premium on default probability is too high to be reasonably explained by most standard models. A possible explanation for the high values of risk neutral PD when LGD is jointly estimated is that a large proportion of the risk premium should be attributed to contagion risk. The firms selected for this study (firms listed in CDX indices) are very large firms whose default is strongly associated with bad times for the market overall, and could cause a contemporaneous negative drop in the market portfolio; see Bai et al. (2015) for a model with this feature. If the study focused on smaller, systemically less important firms, the risk premia may be smaller, though it is possible that the objective default probabilities may be larger in the case of a smaller firm.

Regarding the implications for risk premia associated with LGD, as mentioned above, the estimates of loss given default are much lower than the historical rate of 0.60, which leads to questioning if the risk premium on LGD is negative. A negative risk premium for LGD is difficult to believe, as it implies that when the market is in an overall bad state, default losses become less severe). This would be at odds with financial theories and empirical evidence documenting decreased value of firm collateral and increased fire-sales during times of financial distress, which cause lower recovery rates than usual. However, estimates of LGD being around 0.12 is not inconsistent with a positive risk premium for LGD, because both the estimated risk neutral LGD in this paper, and the historical average 0.60 are not estimates of the unconditional mean of LGD under the two different probability measures. Under the objective measure, LGD is most likely not independent of default probabilities. Instead, low LGD is probably associated with very low PD and vice versa. Then, realized LGD in the historical data is more often drawn from the region of heavy losses. Furthermore, the true conditional objective LGD could be very low in this sample – less than 0.12, which would imply a positive risk premium for LGD. These claims cannot be substantiated in this framework, but they can not be ruled out either, and are a more plausible interpretation of the risk premium for LGD.

5.7 Firm characteristics and credit rating effects

In this section, I study whether firm characteristics are related to estimated levels of LGD and PD using a two-equation panel regression. The LGD and PD measures, which are the two dependent variables in this regression, are the CUGMM point estimate for loss given default L_{cue}, and the average (across maturities) quarterly forward default probability corresponding to L_{cue}, denoted as $avgPD(L_{cue}) \equiv \sum_{j=1}^{40} q_j(\hat{\boldsymbol{\beta}}(L_{cue}))$. I conduct this analysis on the full set of firm-months with non-empty confidence sets. However, for the largest confidence sets, there is very little variation in the CUGMM point estimate for LGD, which is usually very close to 1. These data points may act as highly influential "outliers", or cause a nonlinear effect in the data that cannot be accounted for using the (generalized) linear regression. Thus, I also conduct the panel regression for the subsets of firm-months where the LGD confidence set is less than 0.7, 0.5, 0.2, or 0.1 in length to see whether results change. For each of these five data subsamples, I estimate the LGD and PD equations jointly by stacking the equations. Since both dependent variables take values between 0 and 1, I conduct a fractional response general linear model regression with a logit link function.

The monthly measures of firm characteristics included in the analysis are leverage (ratio of debt to book equity), realized variance of the firm equity price (a proxy for the volatility of firm value), book-to-market ratio (B/M), firm size (log of market capitalization), and credit rating. A constant is included for each equation, and following Papke and Wooldridge (2008), I also add time averages (computed per firm) of each of the explanatory variables to control for firm-level fixed effects.[36] Monthly realized volatility is computed as the square root of the sum of daily "close-to-close" squared returns, using price data from CRSP. All other data used to construct the explanatory variables is sourced from Compustat. Quarterly data items (such as those used to compute leverage and B/M) were assigned to the 3 calendar months spanning the quarter. Monthly credit

[36] The time-series dimension is 146 in this setting, so firm-level fixed effects could probably be estimated directly by including a dummy variable for each group (and omitting the constant), however this involves adding a large number of constants, so I use the approach of Papke and Wooldridge (2008) which was developed for large-N, small-T panels.

ratings are formed by taking the average of Moody's and Standard and Poor's credit ratings, after first converting them to an ordinal number scale, where 1 is the highest rating (Aaa and AAA, respectively). Data for privately held firms is not available on Compustat or CRSP, and thus those firm-months were dropped from the analysis.

Firm leverage and firm value volatility are included in the regression because they are the two sole determinants of credit spreads in the early structural credit model of Merton (1974). In addition, previous studies have found that loss given default is lower for firms in industries that naturally possess more tangible assets that can be sold in the event of a bankruptcy, and firm leverage could reflect this information. Book-to-market is included to see whether there is a relationship between the market's valuation of a firm (relative to it's accounting value) and the market's beliefs about the firm's credit risk.

The inclusion of credit rating as an explanatory variable may seem natural but the expected relationships with the dependent variables are not completely straightforward. Credit ratings are a qualitative measure of how likely it is that a firm will honor its debt obligations, implying a close relationship with objective probability of default, but it is unclear if credit rating is related to the risk neutral counterpart. Furthermore, the expected relationship between credit ratings and LGD is not clear, because firms are not rated by how much they will fall short of their obligations if they do default. However, if firms with higher default probabilities also have lower recovery rates after default, we would observe a negative relationship between LGD and credit quality.

Market capitalization is included to see if there is a relationship between size of a firm and its default probability. Again, the expected relationship is not clear. On one hand, we might expect smaller firms to have more credit risk, in the same manner that "small" firms are shown to be riskier in the Fama-French three-factor model. However, if large firms are more systemically important than small firms, then the risk-premium on default probabilities, and thus risk-neutral default probability would be higher. Ideally, we would want to study the relationships between firm size and the objective default probability and firm size and risk premium separately (and on a larger and more varied cross-section of firms), but unfortunately, this model setup does not allow for any direct measures of risk premium.

The regression results are shown in Table 6. Credit rating has a significantly positive (0.05-level) relationship with PD in the full sample and largest subsample, meaning that as credit rating gets worse (higher ordinal rating), the estimated average PD term structure increases. I also find the same relationship between credit ratings and LGD: firms with worse credit ratings were found to have higher LGD estimates. This result holds in the 4 nested subsamples that drop the largest confidence sets, but is insignificant in the full sample. Months with higher ex-post stock volatility, as measured by realized volatility, see significantly higher default probabilities, but are not related to the point estimate of LGD.

Across all subsamples, leverage ratio is only significant for the PD equation in the full sample, where higher leverage is correlated with higher PD estimates, which is the expected direction of the relationship. Interestingly, book-to-market ratio is positively correlated with LGD in the full sample and two larger subsamples, so that firms that are "undervalued" have higher risk neutral LGD. Also, book-to-market ratio is positively and significantly correlated to PD in the 3 smallest subsamples, though for the full sample, the coefficient on book-to-market ratio is significantly negative, though the magnitude is small.

Finally, market capitalization was not significant for LGD, but all five regressions showed that firm size is negatively related to the default probabilities. This result is analogous to the size premium in equity returns and deserves further study, along with the results for book-to-value ratio.

Table 6: Effects of firm characteristics on estimated LGD and default probability

	Regression sample includes monthly S-sets with length:				
	All	≤ 0.7	≤ 0.5	≤ 0.2	≤ 0.1
			Main regressors:		
Dep. var. = L_{cue}					
Rating	0.02	**0.14***	**0.16***	**0.15***	**0.15***
	(0.05)	**(0.05)**	**(0.04)**	**(0.03)**	**(0.03)**
RVol	0.04	0.10	-0.05	-0.06	-0.07
	(0.06)	(0.11)	(0.08)	(0.07)	(0.08)
D/E	0.25	0.62	0.35	0.63	0.47
	(0.98)	(0.71)	(0.70)	(0.63)	(0.58)
B/M	**0.11***	**0.12***	**0.07***	0.02	0.01
	(0.05)	**(0.04)**	**(0.03)**	(0.03)	(0.03)
MCap	1.98	-1.53	-3.85	-5.04	-3.14
	(4.25)	(6.23)	(8.71)	(5.80)	(4.75)
Dep. var. $= avgPD(L_{cue})$					
Rating	**0.11***	**0.09***	0.04	0.06	0.06
	(0.03)	**(0.04)**	(0.04)	(0.04)	(0.04)
RVol	**0.11**	**0.27***	**0.25***	**0.26***	**0.33***
	(0.06)	**(0.09)**	**(0.08)**	**(0.08)**	**(0.13)**
D/E	**0.68**	0.13	0.33	0.28	0.25
	(0.37)	(0.39)	(0.44)	(0.40)	(0.40)
B/M	**-0.02***	0.01	**0.14***	**0.16***	**0.17***
	(0.01)	(0.01)	**(0.02)**	**(0.02)**	**(0.03)**
MCap	**-17.41***	**-14.34***	**-13.23***	**-13.43***	**-14.16***
	(6.26)	**(5.95)**	**(6.01)**	**(5.58)**	**(5.60)**
No. firm-months	3385	1965	1894	1780	1596

This table presents the coefficient estimates and standard errors for the system of two panel equations in which the dependent variables are the GMM point estimate of LGD (L_{cue}) and the average (across the term structure) default probability corresponding to the LGD point estimate ($avgPD(L_{cue})$). The analysis is conducted using fractional response general linear regression with a logit link function, estimated in Stata with clustered standard errors, which are displayed in brackets. A constant for each equation is also included, along with the time-average (per issuer) of each explanatory variable to control for firm-level fixed effects. Coefficient estimates significant at the 0.05 level are marked with an asterisk, and coefficient estimates significant at the 0.10 level are displayed in bold font.

6 Robustness Checks and Other Extensions

In this Section, I discuss robustness checks and other extensions of the model. Among other changes to the model assumptions, I change the fixed values of λ to allow different amounts of curvature in the Nelson-Siegel function. I also provide a brief discussion of how the model can be applied to a larger set of single-name CDS contracts, or augmented to include simple dynamics.

6.1 Alternate Nelson-Siegel Curves

Results in Section 5 are based on modelling default probability term structure using Nelson-Siegel curves with λ fixed so that the hump of the Nelson-Siegel curvature factor is located at 3.5 years (hereafter NS-3.5, for brevity). To check that results are robust to the choice of λ, I consider two other sets of Nelson-Siegel curves with λ fixed so the hump of the curvature factor is at 2.5 years or 4.5 years, abbreviated NS-2.5 and NS-4.5, respectively. I reproduce the monthly confidence sets for LGD for all thirty corporate issuers with the alternate Nelson-Siegel curves, and Table 7 presents the summaries of confidence set length and locations, comparable to Table 2.

Table 7: Summary of the length and location of Confidence Sets, different Nelson-Siegel models (all firms)

	NS-2.5		NS-3.5		NS-4.5	
	proportion (%)	avg C_L center	proportion (%)	avg C_L center	proportion (%)	avg C_L center
Confidence set length:						
≤ 0.1	37.9	0.13	46.7	0.12	51.7	0.09
(0.1,0.2]	6.5	0.23	7.0	0.20	4.9	0.25
(0.2,0.4]	4.8	0.39	4.4	0.43	3.8	0.44
(0.4,0.6]	3.2	0.53	4.2	0.58	4.4	0.58
(0.6,0.8]	60.	0.58	4.9	0.56	5.6	0.56
(0.8,1]	32.6	0.54	28.7	0.53	25.0	0.53
empty	9.1	–	4.0	–	4.6	–
Testing $L_0 = 0.60$:						
$0.60 \in C_L$	41.0		37.1		33.8	
$0.60 \in C_L \vert \ell(C_L) < 0.5$	2.5		2.9		3.2	
$0.60 \in C_L \vert \ell(C_L) < 0.2$	0.1		0.0		0.1	

Notes: This table reflect three sets of results from different Nelson-Siegel parameterizations. This table groups the monthly confidence sets in bins by their length, and presents the proportion belonging to each bin and the average confidence set center per bin. The bottom rows summarize the proportion of confidence sets that include 0.60 for all firm-months and firm-months whose confidence sets have length less than 0.5 or 0.2.

I find very little change in the results. The estimated confidence sets under NS-2.5 are centered at almost the same level as the NS-3.5 results, while the confidence sets under NS-4.5 are centered slightly lower than the NS-3.5 results. However, all three sets of confidence sets are centered at low values relative to the entire (0,1] parameter space, and the value 0.60 is consistently not included in the firm-months where LGD is precisely estimated. Further, the model is rejected much more often (i.e. LGD confidence set is empty) under NS-2.5 than under NS-3.5. The number of empty sets under NS-4.5 is comparable, but still slightly higher compared to the results under NS-3.5. Thus, the NS-3.5 base parameterization used in Section 5 appears to be the best choice overall for modelling default probability term structure.

6.2 Alternate model assumptions

In addition to changing the curvature of the Nelson-Siegel curves, I consider different durations for the constant DGP assumption, or in other words, I estimate the model semi-monthly and quarterly, instead of monthly, and check whether results change. I estimate these two models on a subset of the data, and the values of LGD in the confidence sets and the distribution of confidence set sizes are similar to the base case. The general advantage of estimating the model over a longer period, is that more data is used, which could allow for more precise estimates of LGD. However, the tradeoff is that the constant DGP assumption must be maintained for a longer period, and the CDS data may not fit as well over the entire period.

The overall results of the base model show that PD and LGD are well identified for some firm-months, but not for all. An important question is whether separate identification of LGD and PD is possible if the term-structure restrictions are loosened. Thus, for a subset of the firm-months, I estimate an extension of the base model that allows LGD to have a linear term structure, $L_{j,t} = \phi_t + \gamma_t j$. I find that the confidence sets for the LGD model, i.e. the joint confidence set for (ϕ_t, γ_t), span large areas of the parameter space, even for firm-months for which the LGD is well-identified in the base model. See section A.2 in the Appendix for some plots of confidence sets.

This extended linear LGD model is not meant as a check on model robustness, but is intended to demonstrate that when an additional degree of modelling freedom is allowed in the LGD term structure, the augmented model is not well-identified. This highlights the general identification problem when estimating PD and LGD from credit spreads, but it also suggests that the base model in this paper allows "maximal" modelling richness, in some sense, while still maintaining enough structural assumptions on the term structure of PD and LGD to guarantee (set or weak) identification.

6.3 Further data and model extensions

For approximately a third of firm-months in this paper, estimated confidence sets for LGD are very large, and thus convey limited information about the true value of risk-neutral expected LGD. To possibly obtain more precise estimates of LGD in these months, the simplest extension is to use a larger set of data, such as spread data from CDS of multiple seniorities.

So far in this paper, I have restricted the data used to credit default swaps on senior unsecured debt (abbreviated as SNRFOR by Markit) because these contracts are most widely available across reference entities, and more heavily traded than other debt tiers: senior secured (SECDOM) or subordinated (SUBLT2). However, the CDS model can be estimated using prices from CDS on debt of multiple seniorities. Generally speaking, these CDS contracts are priced with the same default probability term structure, but with different, though ordered, levels of LGD: $L_{SECDOM} \leq L_{SNRFOR} \leq L_{SUBLT2}$. Then, say for some firm, SUBLT2 CDS prices are available for 8 maturity points from 6 months to 10 years. At the cost of estimating one extra parameter L_{SUBLT2}, we can add this extra data to the estimation and evaluate the profile S-function with both L_{SNRFOR} and L_{SUBLT2} concentrated out. Limitations of this approach are that the data for non-SNRFOR tier CDS are limited, and may be less reliable than senior unsecured CDS prices.

In lieu of using different seniorities of CDS, for which data is very thin, it is also possible to consider CDS contracts with multiple restructuring clauses: ex-restructuring (XR), cum-restructuring (CR), modified restructuring (MR), mod-mod restructuring (MM). For a given reference entity, CDS contracts of the same maturity and debt seniority have the same loss given default, but different default probability term structures. For each point along the term structure of default probabilities, $q_{XR} \leq q_{MR} \leq q_{MM} \leq q_{CR}$. Thus, using, for example, MR contracts in addition to XR contracts requires estimating an additional default probability term structures, so this data extension may not produce more precise estimates of LGD since the parameter dimension increases significantly.

Finally, the model can, in theory, be extended to include simple dynamics as in Diebold and Li (2006), who produce a dynamic Nelson-Siegel model for forecasting daily U.S. yield curves. However, the computational burden to estimate the model will be very significant. In the yield curve model, the 3-dimensional $\boldsymbol{\beta}$ parameters follow a VAR(1) process. In this CDS model, a VAR or GARCH model for $(\boldsymbol{\beta}, L)$ may be suitable, and the data aggregation period could be shortened, from one calendar month as in Section 5 to maybe 5 days. This dynamic approach is not pursued in this study, because the focus is on producing robust estimates of LGD that do not depend time series modelling of credit default dynamics.

7 Conclusion

This paper combines a term structure model of credit default swaps (CDS) with weak-identification robust methods to jointly estimate the probability of default and the loss given default of the underlying firm. In general, identification is bought using assumptions, and existing models that jointly estimate loss given default and probability of default impose parametric time series specifications for stochastic default intensity whose assumptions are unverifiable in the data, and for which there is no consensus of best specification.

Instead, this paper forgoes the parametric time series restrictions, and only models the term structure of default probability and loss given default at a point in time. As a result of weakening the assumptions, the model is not globally identified, but valid inference is conducted using econometric methods robust to weak identification.

As an empirical contribution, I estimate the model independently each month on 30 U.S. firms for a period spanning 146 months, and construct 95% confidence sets for loss given default. The results show that informative (small) confidence sets centered close to 0.10 are estimated for half of the firm-months in the sample. For almost all of these firm-months, the conventional loss given default value of 0.60 is rejected. In addition, risk-neutral default probabilities, and hence risk premia on default probabilities, are underestimated when loss given default is exogenously fixed at the conventional value instead of estimated from the data.

References

Andrews, D. W. and X. Cheng (2012). Estimation and inference with weak, semi-strong, and strong identification. *Econometrica 80*(5), 2153–2211.

Annaert, J., A. G. Claes, M. J. De Ceuster, and H. Zhang (2013). Estimating the spot rate curve using the nelson–siegel model: A ridge regression approach. *International Review of Economics & Finance 27*, 482–496.

Bai, J., P. Collin-Dufresne, R. S. Goldstein, and J. Helwege (2015). On bounding credit event risk premia. *The Review of Financial Studies*.

Barclays Capital (2010). Standard corporate CDS handbook.

Berndt, A., R. Douglas, D. Duffie, M. Ferguson, and D. Schranz (2008). Measuring default risk premia from default swap rates and EDFs. *Tepper School of Business Working Paper*, 49.

Berndt, A., R. A. Jarrow, and C. Kang (2007). Restructuring risk in credit default swaps: An empirical analysis. *Stochastic Processes and their applications 117*(11), 1724–1749.

Canova, F. and L. Sala (2009). Back to square one: identification issues in dsge models. *Journal of Monetary Economics 56*(4), 431–449.

Christensen, J. H. E. (2005). Joint estimation of default and recovery risk: A simulation study. Technical report.

Christensen, J. H. E. (2007). Joint default and recovery risk estimation: an application to CDS data. Technical report.

Conrad, J., R. F. Dittmar, and A. Hameed (2013). Cross-market and cross-firm effects in implied default probabilities and recovery values. *Working Paper*.

Das, S. R. and P. Hanouna (2009). Implied recovery. *Journal of Economic Dynamics and Control 33*(11), 1837–1857.

Diebold, F. X. and C. Li (2006). Forecasting the term structure of government bond yields. *Journal of econometrics 130*(2), 337–364.

Driessen, J. (2005). Is default event risk priced in corporate bonds? *Review of Financial Studies 18*(1), 165–195.

Duffie, D. (1998). Defaultable term structure models with fractional recovery of par. *Graduate School of Business, Stanford University*.

Duffie, D. and K. J. Singleton (1999). Modeling term structures of defaultable bonds. *Review of Financial studies 12*(4), 687–720.

Elkamhi, R., K. Jacobs, and X. Pan (2010). The cross-section of recovery rates and default probabilities implied by credit default swap spreads. *Journal of Financial and Quantitative Analysis*, 1–51.

Guha, R. and A. Sbuelz (2005). Structural recovery of face value at default. *SSRN Working Paper*.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 1029–1054.

Houweling, P. and T. Vorst (2005). Pricing default swaps: Empirical evidence. *Journal of International Money and Finance 24*(8), 1200–1225.

Le, A. (2014). Separating the components of default risk: A derivative-based approach. *Quarterly Journal of Finance*.

Longstaff, F. A., S. Mithal, and E. Neis (2005). Corporate yield spreads: Default risk or liquidity? new evidence from the credit default swap market. *The Journal of Finance 60*(5), 2213–2253.

Madan, D. B., G. Bakshi, and F. X. Zhang (2006). Understanding the role of recovery in default risk models: Empirical comparisons and implied recovery rates. *Working Paper*.

Markit Group Ltd. (2010). Markit.com user guide: CDS & bonds.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates*. *The Journal of Finance 29*(2), 449–470.

Nason, J. M. and G. W. Smith (2008). Identifying the new keynesian phillips curve. *Journal of Applied Econometrics 23*(5), 525–551.

Nelson, C. R. and A. F. Siegel (1987). Parsimonious modeling of yield curves. *Journal of business*, 473–489.

Pan, J. and K. J. Singleton (2008). Default and recovery implicit in the term structure of sovereign CDS spreads. *The Journal of Finance 63*(5), 2345–2384.

Papke, L. E. and J. M. Wooldridge (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics 145*(1), 121–133.

Schläfer, T. and M. Uhrig-Homburg (2014). Is recovery risk priced? *Journal of Banking & Finance 40*, 257–270.

Schneider, P., L. Sögner, and T. Veža (2011, 1). The economic role of jumps and recovery rates in the market for corporate default risk. *Journal of Financial and Quantitative Analysis 45*, 1517–1547.

Song, J. (2007). Loss given default implied by cross-sectional no arbitrage. *Working Paper*.

Stock, J. H. and J. H. Wright (2000). GMM with weak identification. *Econometrica 68*(5), 1055–1096.

A Appendix for Paper

A.1 Additional Tables and Figures

Table A.1: CDS Reference Entities and their Credit Ratings

Markit ticker	reference entity	sector	ratings range
	Investment Grade (listed in CDX.NA.IG.17)		
AA	Alcoa Inc.	basic materials	A- ; BBB-
CAG	ConAgra Foods, Inc.	consumer goods	BBB+ ; BBB-
MO	Altria Group, Inc.	consumer goods	BBB+ ; BBB
WHR	Whirlpool Corporation	consumer goods	BBB+ ; BBB-
AZO	Autozone, Inc.	consumer services	BBB+ ; BBB
SWY	Safeway Inc.	consumer services	BBB ; B
YUM	YUM! Brands, Inc.	consumer services	BBB ; BB
AXP	American Express Company	financials	A+ ; BBB+
GEcc	General Electric Capital Corp.	financials	AAA ; AA+
ARW	Arrow Electronics, Inc.	industrials	BBB- ; BBB-
CSX	CSX Corporation	industrials	BBB+ ; BBB-
UNP	Union Pacific Corp.	industrials	A ; BBB-
UNH	Unitedhealth Group Inc.	healthcare	A+ ; A-
APC	Anadarko Petroleum Corp.	energy/oil&gas	BBB+ ; BBB-
COP	Conoco Phillips	energy/oil&gas	A ; A-
VLOC	Valero Energy Corp.	energy/oil&gas	BBB ; BBB-
CSC	Computer Sciences Corp.	technology	A ; BBB
XRX	Xerox Corp.	technology	BBB ; BB-
COX	COX Communications, Inc.	telecommunications	BBB+ ; BBB-
AEP	American Electric Power Company, Inc.	utilities	BBB ; BBB
	High Yield (listed in CDX.NA.HY.17)		
WY	Weyerhaeuser Company	basic materials	BBB ; BBB-
GT	The Goodyear Tire & Rubber Company	consumer goods	BB ; B+
LTD	L Brands, Inc.	consumer services	BBB ; BB
RCL	Royal Caribbean Cruises Ltd.	consumer services	BBB- ; BB-
RSH	Radioshack Corp.	consumer services	A- ; D
TSG	Sabre Holdings Corp.	consumer services	not avail.
CMS	CMS Energy Corp.	energy/oil&gas	BBB+; BB
MGIC	MGIC Investment Corp.	financials	A ; CCC
THC	Tenet Healthcare Corp.	healthcare	B+ ; B
AMD	Advanced Micro Devices, Inc.	technology	BB- ; CCC+

This table lists the CDS reference entities studied in the paper, their sector, and the range of Standard & Poor's long-term bond ratings for the sample period from January 2004 through September 2015. The first column displays Markit tickers for the reference entities, though the full Markit ticker for General Electric Capital Corp. is GE-CapCorp, and the full Markit ticker for COX Communications Inc. is COX-CommInc. The sector classifications are as reported in the Markit CDX index annexes.

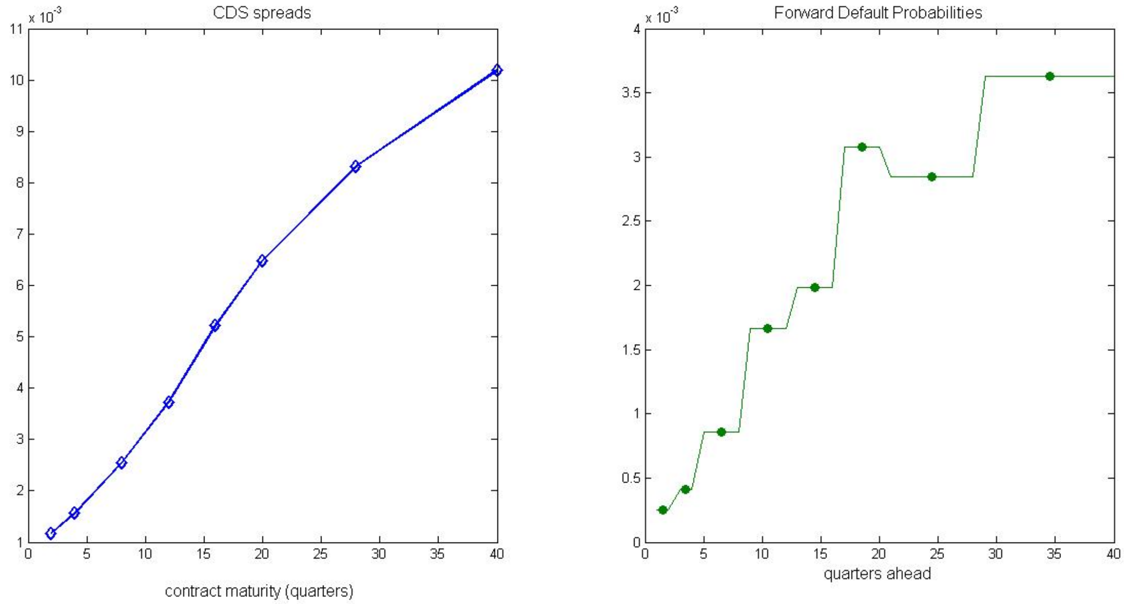


Figure A.1: *The left panel of this figure presents the CDS spread term structure in dollars on a sample day for a sample firm (Xerox on Nov. 21, 2007). The right panel presents the 8-step forward default probability curve stripped from the CDS term structure with LGD is fixed at 0.40.*

Table A.2: Summary of lengths and locations of 90% Confidence Sets ($\alpha = 0.10$)

	full sample		Jan04–Dec06		Jan07–Dec09		Jan10–Dec12		Jan13–Feb16	
	Prop. (%)	avg C_L center	Prop. (%)	avg C_L center	Prop. (%)	avg C_L center	Prop. (%)	avg C_L center	Prop. (%)	avg C_L center
confidence set length:										
≤ 0.1	52	0.11	50	0.08	22	0.10	60	0.11	78	0.14
(0.1,0.2]	5	0.24	5	0.15	6	0.23	5	0.34	4	0.22
(0.2,0.4]	5	0.47	4	0.39	5	0.41	7	0.56	3	0.42
(0.4,0.6]	3	0.61	3	0.58	3	0.59	6	0.65	1	0.55
(0.6,0.8]	5	0.57	5	0.52	7	0.57	4	0.58	3	0.61
(0.8,1]	25	0.53	27	0.52	49	0.52	16	0.53	7	0.55
empty	5	–	5	–	8	–	2	–	5	–
testing $L_0 = 0.60$:										
$0.60 \in C_L$	33		36		59		26		11	
$0.60 \in C_L \vert \ell(C_L) < 0.5$	3		3		5		6		1	
$0.60 \in C_L \vert \ell(C_L) < 0.2$	0		0		0		0		0	

This table shows results based on asymptotic 90% confidence sets for LGD. The confidence set sizes and locations, and the proportion including 0.60 are presented. For comparison, this table replicates the summaries of the 95% confidence sets found in tables 2 and 3.

A.2 Linear LGD term structure

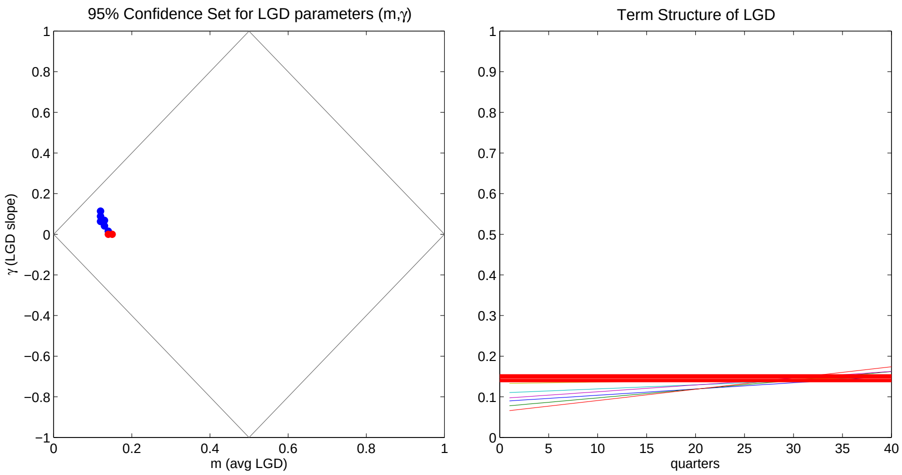
This section presents results of LGD estimation for the extended model allowing LGD to have a linear term structure: $L_{j,t} = \phi_t + \gamma_t j$, for $j = 1, \ldots, 40$ quarters. The PD term structure remains modelled by the three-parameter Nelson-Siegel model as in Equation (5).

For the purposes of plotting the confidence sets for LGD, we reparameterize the LGD term structure to $L_{j,t} = m_t + \gamma_t(j - 20.5)$, so that m_t is the mean level of the LGD term structure over quarters 1 through 40 (in the t-th month).

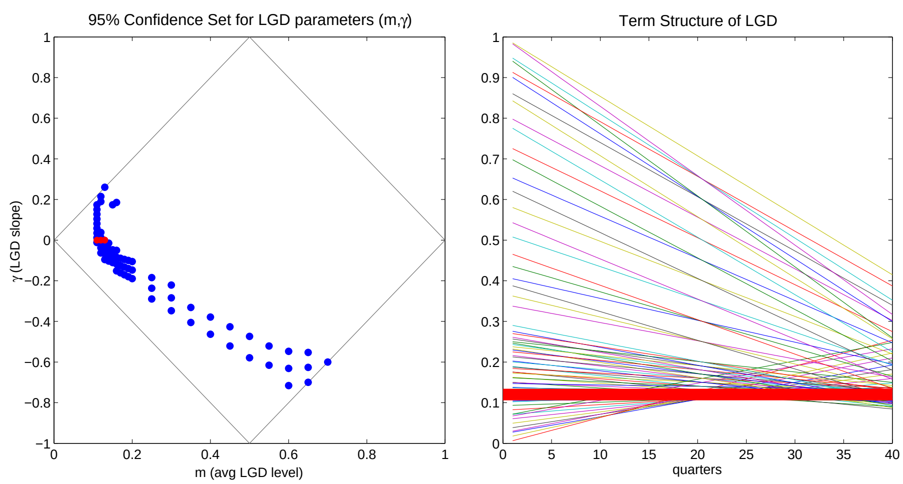
The next six plots show the results of LGD estimates for three firm-months, LTD in April 2011, AA in Dec. 2010, and XRX in July 2007. The three plots on the left show the (roughly plotted) confidence sets for (m_t, γ_t) for the three firm-months. The possible parameter space for (m_t, γ_t) is the inside of the diamond in the left plots. Parameter values outside of this diamond would cause the LGD term structure to take values outside of (0,1] for some quarters. The plots on the right show the estimated term structure of LGD over quarters 1 through 40. The focus is on the estimates from the extended linear LGD model, however, the estimates from the base model with flat LGD are also plotted in red in all plots for comparison and context.

The two plots in the top row show that LGD confidence set is very small for both the linear LGD model and flat LGD model. However, results as these are very rare – less than 1% of the firm-months estimated. Much more common are results like those in the middle and bottom rows. In the middle row, when LGD is contrained to be flat, the LGD confidence set is very small, however, the confidence set under the linear LGD model takes up a larger portion of the parameter space, and when translated to the term structure plot on the middle-right, we clearly see that a wide variety of LGD term structures with different levels and slopes are contained in the confidence set. The bottom row shows the same lack of identification of LGD, but even more extreme.

L Brands (LTD) Apr 2011



Alcoa (AA) Dec 2010



Xerox (XRX) Jul 2007