Internet Appendix for:
"Cyclical Dispersion in Expected Defaults"

João F. Gomes[*] Marco Grotteria[†] Jessica Wachter [‡]

August, 2017

Contents

[*]The Wharton School, University of Pennsylvania. Email: gomesj@wharton.upenn.edu

[†]The Wharton School, University of Pennsylvania. Email: grottm@wharton.upenn.edu

[‡]The Wharton School, University of Pennsylvania and NBER. Email: jwachter@wharton.upenn.edu

1 Robustness Tests

The results presented in the main text are robust to the inclusion of other variables. The results are also robust to the choice of frequency. In this latter case, portfolios are constructed directly using Compustat annual data. The following empirical results are run on non-financial non-regulated companies only.

1.1 Multivariable Forecasting of Macroeconomic Quantities

Table 1 presents the results from a host of additional robustness tests for the OLS predictability regression of output (Panel A) and investment growth (Panel B) at 1 quarter and 2 quarters. As independent variables, we use other economic series that have been empirically found helpful predictors of the economic cycle.

In column (a) we forecast macroeconomic quantities using only the average Expected Default Frequency of net repayers. The estimate exhibits a negative sign and is statistically significant at 1% level in all the cases. In column (b) we regress the same variables onto *Dispersion*, as shown in the main text. In column (c) we forecast the macroeconomic quantities using *Dispersion* and the Gilchrist and Zakrajšek (2012) Excess Bond Premium. Results are impaired by the high collinearity between the two series. *Dispersion*, which is a naive linear combination of the EDF of repayers and the one of issuers and not the one maximally correlated with EBP, has a correlation of 0.65 with EBP. In column (d) we regress the macroeconomic series onto *Dispersion*, the log of the price-dividend ratio, the term spread and the lagged dependent variable. *Dispersion* remains statistically and economically significant.

1.2 Annual Portfolios

In this section, we present the OLS regression estimates from predicting macroeconomic quantities and bond returns using both our *Dispersion* measure and Greenwood and Hanson (2013) *ISSEDF*. Results align with the quarterly estimates presented in the main text. Please notice that Greenwood and Hanson measure is a decile-based measure and not a probability measure, which requires a different interpretation of the coefficient estimates compared to ours. Furthermore, please notice that Greenwood and Hanson subtract the average EDF decile of repayers from the average EDF decile of issuers, while we subtract the average EDF of issuers from the one of repayers (as repayers have always a higher average EDF). The interpretation with their measure is as follows: when firms with high net debt issuance have EDFs that are on average one decile higher than firms with low net debt issuance, excess returns on high yield bonds is expected to be 12% lower next year and excess returns on investment grade bonds 2% lower.

2 Additional Empirical Results

This section reports a few additional results concerning *Dispersion* in credit quality and other dispersion measures based on individual firms' credit-risk.

Figure 1 plots both the CBOE VIX and our main dispersion measure (based on debt repayers and issuers). Dispersion is closely related to the CBOE Volatility Index but there are some periods, like the early 90s or 2015, where a higher volatility did not necessarily imply higher default risk and vice-versa.

Table 4 and 5 report the predictability regressions based on the dispersion in EDFs between firms with smallest asset growth and firms with the highest asset growth in a given quarter.

Figure 2 shows the average distress risk measure for repayers and issuers computed from the estimates by Campbell, Hilscher, and Szilagyi (2008) (Table IV, 0 lag). In table 6 and 7, we reproduce the results for the predictability regressions when dispersion is based on Campbell et al. (2008) distress measure.

Table 8 demonstrates that *Dispersion* in credit quality (the one based on debt repayment) is also a strong predictor of unemployment growth.

Table 9 reports the results from out-of-sample predictability of bond returns. To interpret the economic significance of such results we use the formula suggested by Cochrane (1999). It can be proven that the Sharpe ratio (s^*) earned by an investor who uses the entire information from those predictability regressions (R^2) and the Sharpe ratio (s_0) otherwise earned via a buy-and-hold strategy are related according to the following formula

$$s^* = \sqrt{\frac{s_0^2 + R^2}{1 - R^2}}$$

Given an annualized Sharpe ratio for the buy-and-hold strategy of 0.375 (obtained multipliying the quarterly Sharpe ratio by $\sqrt{2}$) and a predictive quarterly R^2 of 11.7%, the implied annualized Sharpe ratio for an active investor is about 0.651. As regards investment-grade bonds, the annualized Sharpe ratio for a buy and hold strategy equals 0.54 and once we account for the predictive information available to investors, we observe that an active investor could reach a Sharpe ratio of 0.729. In the same table we report the 25–75 confidence interval of the R^2 statistics based on 1000 bootstrapped samples of the same size of the actual sample used for the estimation.

References

Campbell, J. Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. Journal of Finance 63, 2899–2939.

Cochrane, J. H., 1999. New facts in finance. Economic Perspectives, Federal Reserve Bank of Chicago 23, 36–58.

Gilchrist, S., Zakrajšek, E., 2012. Credit spreads and business cycle fluctuations. American Economic Review 102, 1692–1720.

Greenwood, R., Hanson, S. G., 2013. Issuer quality and corporate bond returns. Review of Financial Studies 26, 1483–1525.

Newey, W. K., West, K. D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.

Table 1. Multivariable forecasting - Horizon k quarters

Panel A: $\overline{\Delta\,\text{GDP}}_{t\to t+k}$

	k =1				k =2			
	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(d)
EDF_t^R	-0.309^{***}				-0.236^{***}			
	[−5.26]				[−4.18]			
Dispersion_t		-0.412^{***}	-0.233^{**}	-0.318^{***}		-0.323^{***}	-0.164	-0.255^{***}
		[−6.42]	[−2.16]	[−4.26]		[−4.69]	[−1.38]	[−3.01]
EBP_t			-0.342^{**}				-0.302^{*}	
			[−2.24]				[−1.81]	
$\log(\text{pd})_t$				0.001				0.0002
				[0.45]				[0.14]
TS_t				0.481^{**}				0.647^{***}
				[2.09]				[3.67]
$\overline{\Delta\,\text{GDP}}_{t-k\to t}$				0.242^{**}				0.230^{**}
				[2.52]				[1.99]
$\mathbf{R^2}$	0.1181	0.1128	0.1424	0.2116	0.1006	0.0999	0.1336	0.2490

Panel B: $\overline{\Delta\,\mathrm{I}}_{t\to t+k}$

	k =1				k =2			
	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(d)
EDF_t^R	-1.330^{***}				-1.003^{***}			
	[−6.60]				[−4.39]			
$\mathrm{Dispersion}_t$		-1.740^{***}	-0.964^{**}	-1.331^{***}		-1.343^{***}	-0.631	-1.221^{***}
		[−6.53]	[−2.48]	[−4.39]		[−4.45]	[−1.52]	[−3.49]
EBP_t			-1.474^{***}				-1.354^{**}	
			[−2.95]				[−2.24]	
$\log(\mathrm{pd})_t$				0.006				0.005
				[1.17]				[1.04]
TS_t				2.034^{**}				2.758^{***}
				[2.27]				[4.42]
$\overline{\Delta\,\mathrm{I}}_{t-k\to t}$				0.295^{***}				0.172
				[2.99]				[1.30]
$\mathbf{R^2}$	0.1689	0.1548	0.1972	0.3138	0.1323	0.1262	0.1758	0.3045

Significance levels : $*$: 10% $**$: 5% $***$: 1%

Notes: The table reports coefficients and R^2 statistics from predictive regressions of average GDP growth (Panel A) and average Investment Growth (Panel B) over one and two quarterly horizons for four different specifications. We define dispersion as average EDF of repayers minus average EDF of issuers. EDF_t^R is the average expected default frequency of repayers only. EBP is the quarterly average of the monthly series of Gilchrist and Zakrajšek (2012) excess bond premium. log(pd) is the log of the price-dividend ratio of the CRSP index (all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ). TS refers to the term spread and is computed as the yield (at a quarterly level) on Treasury nominal securities of 10 year "constant maturity" minus the yield (at a quarterly level) on the 3-Month Treasury Bill. We construct t-statistics from Newey and West (1987) standard errors, with $k-1$ lags, where k is the regression horizon. Data are quarterly from January 1976 until September 2013. Statistical significance levels at 10%, 5% and 1% are denoted by *, ** and ***, respectively.

Table 2. Forecasting Economic Activity: Horizon 1 year

Panel A: Δ Per Capita GDP$_{t \to t+1}$

	(a)	(b)
β	-0.35^{**} [−2.23]	-0.14 [−0.22]
$\mathbf{R^2}$	0.1011	0.0009

Panel B: Δ Per Capita Investment$_{t \to t+1}$

	(a)	(b)
β	-1.38^{**} [−2.43]	-3.13 [−1.37]
$\mathbf{R^2}$	0.1149	0.0324

Significance levels : $*$: 10% $**$: 5% $***$: 1%

Source: Bureau of Economic Analysis, CRSP/Compustat merged, CRSP

Notes: The table presents OLS coefficient estimates and t-statistics in parentheses. Standard errors are White (1980) robust standard errors. Columns (a) presents the results for the raw measure of *Dispersion* in credit quality. Column (b) is Greenwood and Hanson (2013) ISSEDF. The frequency is annual. Data are from 1968 to 2008.

Table 3. Forecasting Excess Returns on Bonds ($rx_{t \to t+1}$): Horizon 1 year

Panel A: Investment Grade

	(a)	(b)
β	0.71	-1.89
	[1.36]	[-1.04]
$\mathbf{R^2}$	0.0482	0.0184

Panel B: High Yield

	(a)	(b)
β	2.93^{*}	-12.03^{**}
	[1.83]	[-2.35]
$\mathbf{R^2}$	0.1362	0.1131

Significance levels : $*$: 10% $**$: 5% $***$: 1%

Source: Barclays Capital, Global Financial Data, CRSP/Compustat merged, CRSP

Notes: The table presents OLS coefficient estimates and t-statistics in parentheses. Standard errors are White (1980) robust standard errors. Columns (a) presents the results for the raw measure of *Dispersion* in credit quality. Column (b) is Greenwood and Hanson (2013) ISSEDF. The frequency is annual. Investment-grade bond data are from January 1976 until September 2013. High-yield bond data are from January 1987 to June 2013.

Table 4. Forecasting Macroeconomic Quantities: Dispersion based on asset growth

	Horizon k				
	1	2	3	4	8
			Panel A: GDP		
β	-0.39^{***}	-0.28^{***}	-0.22^{***}	-0.17^{***}	-0.08
	[−6.24]	[−4.47]	[−3.52]	[−2.93]	[−1.40]
R^2	0.1205	0.0935	0.0705	0.0527	0.0160
			Panel B: Investment		
β	-1.67^{***}	-1.23^{***}	-0.84^{***}	-0.55^{**}	-0.10
	[−5.56]	[−4.65]	[−3.39]	[−2.55]	[−0.52]
R^2	0.1728	0.1299	0.0748	0.0388	0.0022

Source: Bureau of Economic Analysis, CRSP/Compustat merged, CRSP

Notes: Estimation of

$$\overline{\Delta y}_{t\to t+k} = \alpha + \beta\,\mathrm{DAG}_t + \epsilon_{t+k}.$$

The table reports coefficients and R^2 statistics from predictive regressions of average GDP (Panel A) and average investment growth (Panel B) over various horizons onto dispersion in EDF values with sorting of individual firms based on asset growth (DAG). Each quarter we sort firms in the data into quintiles based on change in book value of assets divided by lagged total assets. We define dispersion based on asset growth as the average EDF of the bottom asset growth quintile minus the average EDF of the top asset growth quintile. We construct t-statistics from Newey and West (1987) standard errors, with $k-1$ lags, where k is the regression horizon. Data are quarterly from January 1976 until September 2013. Statistical significance levels at 5% and 1% are denoted by ** and ***, respectively.

Table 5. Forecasting Excess Returns on Bonds: Dispersion in asset growth

	Horizon k				
	1	2	3	4	8
			Panel A: Investment Grade		
β	0.79	0.93***	0.83***	0.74***	0.46***
	[1.50]	[2.81]	[2.89]	[3.14]	[2.77]
R^2	0.0455	0.1186	0.1326	0.1375	0.1153
			Panel B: High Yield		
β	2.26**	2.42***	2.13***	1.86***	1.33***
	[2.30]	[3.32]	[3.20]	[3.22]	[3.99]
R^2	0.0960	0.1815	0.2272	0.2430	0.2903

Source: Barclays Capital, Global Financial Data, CRSP/Compustat merged, CRSP

Notes: Estimation of

$$\overline{rx}_{t\to t+k} = \alpha + \beta\,\mathrm{DAG}_t + \epsilon_{t+k}.$$

The table reports coefficients and R^2 statistics from predictive regressions of average excess log returns on bonds over various horizons onto dispersion in EDF values with sorting of individual firms based on asset growth (DAG). Panel A reports results for investment grade bonds; panel B reports results for high yield bonds. Each quarter we sort firms in the data into quintiles based on change in book value of assets divided by lagged total assets. We define dispersion based on asset growth as the average EDF of the bottom asset growth quintile minus the average EDF of the top asset growth quintile. We construct t-statistics from Newey and West (1987) standard errors, with $k-1$ lags, where k is the regression horizon. Investment-grade bond data are quarterly from January 1976 until September 2013. High-yield bond data are quarterly from January 1987 to June 2013. Statistical significance levels at 5% and 1% are denoted by ** and ***, respectively.

Table 6. Forecasting Macroeconomic Quantities: Alternative model

	Horizon k				
	1	2	3	4	8
		Panel A: GDP			
β	-0.42^{***}	-0.27^{**}	-0.19^{*}	-0.12	-0.03
	[−3.79]	[−2.54]	[−1.90]	[−1.25]	[−0.34]
R^2	0.0732	0.0443	0.0265	0.0126	0.0012
		Panel B: Investment			
β	-1.97^{***}	-1.41^{***}	-0.90^{**}	-0.46	0.05
	[−5.11]	[−3.22]	[−2.44]	[−1.44]	[0.18]
R^2	0.1265	0.0878	0.0443	0.0138	0.0003

Source: Bureau of Economic Analysis, CRSP/Compustat merged, CRSP

Notes: Estimation of

$$\overline{\Delta y}_{t\to t+k} = \alpha + \beta\,\text{Distress}_t + \epsilon_{t+k}.$$

The table reports coefficients and R^2 statistics from predictive regressions of average GDP (Panel A) and average investment growth (Panel B) over various horizons onto dispersion in credit quality between debt repayers and issuers, where credit quality is measured using the distress risk measure based on Campbell, Hilscher, and Szilagyi (2008). We construct t-statistics from Newey and West (1987) standard errors, with $k-1$ lags, where k is the regression horizon. Data are quarterly from January 1976 until September 2013. Statistical significance levels at 10%, 5% and 1% are denoted by *, ** and ***, respectively.

Table 7. Forecasting Excess Returns on Bonds: Alternative model

	Horizon k				
	1	2	3	4	8
			Panel A: Investment Grade		
β	0.67	1.23***	1.12***	0.94***	0.60***
	[1.13]	[3.50]	[3.17]	[3.07]	[2.65]
R^2	0.0173	0.1102	0.1257	0.1174	0.1018
			Panel B: High Yield		
β	2.68**	3.39***	2.92***	2.50***	1.85***
	[2.31]	[3.88]	[3.29]	[3.23]	[6.13]
R^2	0.0698	0.1831	0.2186	0.2223	0.2925

Source: Barclays Capital, Global Financial Data, CRSP/Compustat merged, CRSP

Notes: Estimation of

$$\overline{rx}_{t\to t+k} = \alpha + \beta\,\mathrm{Distress}_t + \epsilon_{t+k}.$$

The table reports coefficients and R^2 statistics from predictive regressions of average excess log returns on bonds over various horizons onto dispersion in credit quality between debt repayers and issuers, where credit quality is measured using the distress risk measure based on Campbell, Hilscher, and Szilagyi (2008). Panel A reports results for investment grade bonds; panel B reports results for high yield bonds. We construct t-statistics from Newey and West (1987) standard errors, with $k-1$ lags, where k is the regression horizon. Investment-grade bond data are quarterly from January 1976 until September 2013. High-yield bond data are quarterly from January 1987 to June 2013. Statistical significance levels at 10%, 5% and 1% are denoted by *, ** and ***, respectively.

Table 8. Forecasting Macroeconomic Quantities: Dispersion in debt repayment

	Horizon k				
	1	2	3	4	8
			Unemployment		
β	3.66*** [7.59]	3.08*** [6.21]	2.60*** [5.47]	2.17*** [4.78]	1.10*** [2.69]
R^2	0.2536	0.2142	0.1715	0.1330	0.0490

Source: Bureau of Economic Analysis, CRSP/Compustat merged, CRSP

Notes: Estimation of

$$\overline{\Delta Unemp.}_{t\to t+k} = \alpha + \beta\,\text{Dispersion}_t + \epsilon_{t+k}.$$

The table reports coefficients and R^2 statistics from predictive regressions of average unemployment growth over various horizons onto dispersion in credit quality (*Dispersion*). We define dispersion as average EDF of repayers minus average EDF of issuers. We construct t-statistics from Newey and West (1987) standard errors, with $k-1$ lags, where k is the regression horizon. Data are quarterly from January 1976 until September 2013. Statistical significance levels at 5% and 1% are denoted by ** and ***, respectively.

Table 9. Bond Return Predictions - Horizon 1 Quarter

	High Yield		Investment Grade	
	R^2	CI [25, 75]	R^2	CI [25, 75]
Out-of-sample	0.1169	[0.0330, 0.2245]	0.0950	[0.0269, 0.1197]

Notes: We report out-of-sample percentage R^2 for OLS forecasts of 1-quarter bond excess returns from from January 1976 until September 2013 for investment grade bonds and from January 1987 to June 2013 for High-yield bonds. The predictor variable is *Dispersion* in credit quality. Our out-of-sample procedure splits the sample after the first 55 observations, uses the first 55 observations as a training window, and recursively forecasts returns using all available information to obtain parameter estimates, i.e. using an incrasing estimation window. We also report the 25-75 confidence interval for the R^2 computed using 1000 bootstrapped samples of the same size of the actual samples.

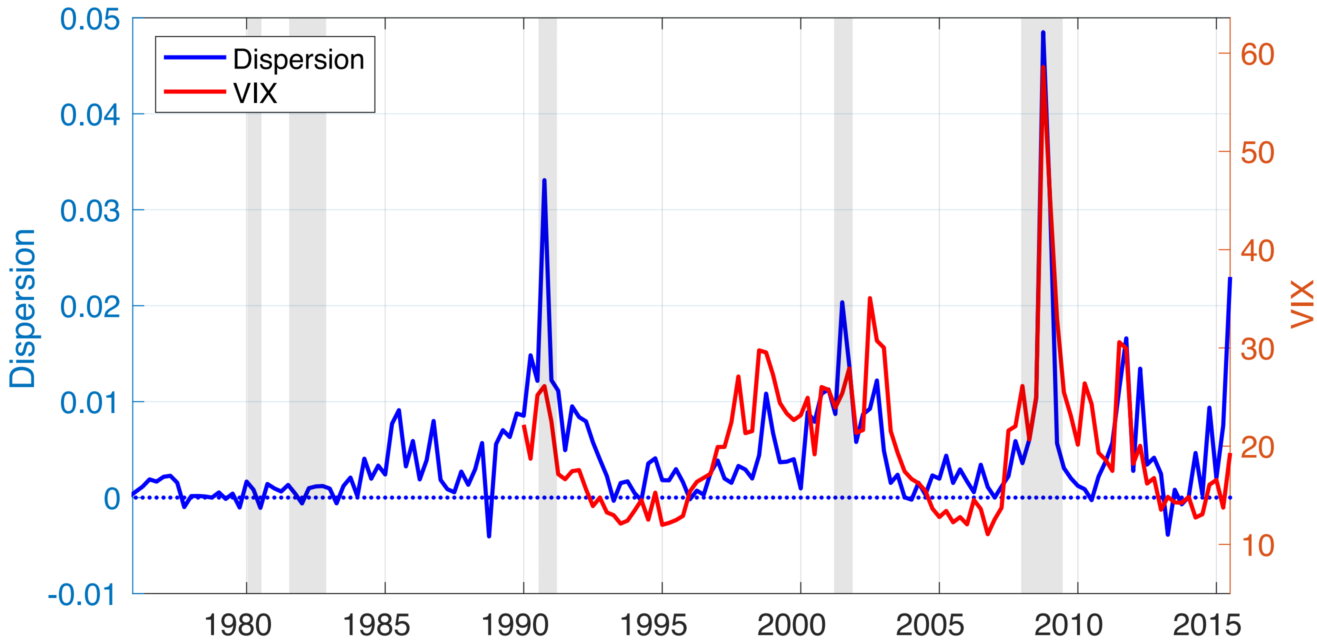

Fig. 1. Dispersion and VIX. The figure shows the dispersion between the average EDF for firms which repay their debt minus the average EDF for issuers and compares the series with the quarterly average of VIX. The shaded areas correspond to NBER recessions.

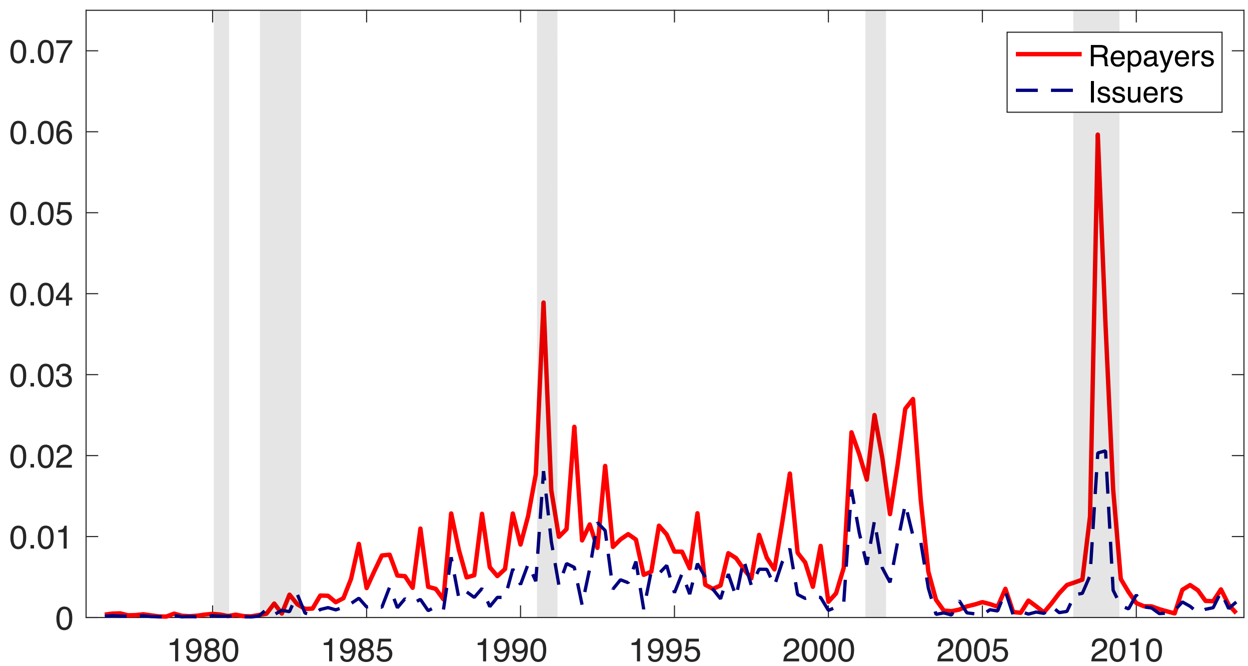

Fig. 2. Distress risk based on Campbell, Hilscher, and Szilagyi (2008). The figure shows the average distress risk based on Campbell, Hilscher, and Szilagyi (2008) of debt repayers and issuers. The shaded areas correspond to NBER recessions.