Appendix A Data

A.1 Definition of default

As described in Section 2, the default definition we adopt is the following: a firm is considered to be in default whenever one of the following events occurs: the firm is declared legally bankrupt; has suspended payments; has negotiated a debt composition settlement; is undergoing a reconstruction; or is distraint without assets. The data on Swedish public and private firms that we use to construct the default variable have been provided by Upplysningscentralen AB (UC), the main Swedish credit bureau, jointly owned by the Swedish banks. UC taps its information from Tingsrätten, the District Courts, Bolagsverket, the Swedish Companies Registration Office (SCRO), and Kronofogdemyndigheten, the Swedish Enforcement Authority, i.e. the institutions that deal with the legal formalities in firms´ bankruptcy processes.[A.1]

UC stores information on firms minor and major distress events in two databases, AM and JP. In the first database, AM, variables are constructed by giving each type of event a label AMTYPXX, a Swedish acronym for remark type, and an integer number suffix. For example, AMTYP12 is a dummy variable indicating if a firm has suspended its payments, or not. The second database, JP, contains 27 variables in total, on various milestones and stages for a broader category of major (mostly, but not exclusively) distress events that may occur for incorporated Swedish firms. From this database we take variables that are related to legal bankruptcy: "bankruptcy procedures started," "bankruptcy procedures concluded," "bankruptcy procedures concluded with a surplus," "bankruptcy procedures continued," and "declared bankrupt." Moreover, we also include: "negotiations on a debt composition settlement started," and "negotiations on a debt composition settlement concluded."

If any of the above distress-event dummy variables equals one at some point in our sample period, the firm in question is considered to be in default in that particular quarter. In the following quarter, we let the firm exit our data set. If more than one of these distress events are observed for a specific firm over our sample period, we assume the firm in question defaulted in the first instance. An additional variable we use from the second data set indicates if a "bankruptcy [was] cancelled" by a court. Over the whole sample period (i.e., in-sample and out-of-sample) this occurs 24 times, and 16 of these 24 events concern firms that subsequently end up in default. We treat firms for which the bankruptcy status was cancelled by the District

[A.1] Currently, i.e., on July 9, 2010, there are 462 publicly listed firms out of roughly 250,000 active, limited liability firms.

Court as healthy until the data indicate otherwise. Moreover, we let firms that default but re-emerge from their default status exit the data set after the quarter in which default takes place; they re-enter in the quarter in which UC registered that the default status had been "removed."

Our decision to let firms that default exit the data set in the subsequent quarter is based on the following statistics: out of 161,550 defaults in the entire data set, 148,874 are terminal in the sense that subsequently no new information on the firms appears in any of the databases.[A.2] The remaining 12,676 observations concern 6,638 firms that default twice within the sample period. Thus, 5,999 firms end up in terminal default at the second occurrence, while 339 re-emerge even after the second default. No firm defaults more than twice in the sample period.

Out of the 148,874 first-time-is-terminal defaults, an overwhelming majority of 145,684 are due to legal bankruptcy declarations. Roughly 60 percent of these firms experience a second default-triggering distress event simultaneously, i.e., in the same quarter. In almost all cases (98 percent) these are due to the event "bankruptcy proceedings initiated." In most (88 percent) of the remaining terminal defaults, i.e., those that are not legal bankruptcies, are associated with "distraint, no assets." The remaining distress events account for less than 1 percent of the first-time-is-terminal defaults.

For the firms that re-emerge after a default, the first default involves a legal bankruptcy in less than 2 percent of all cases and "distraint, no assets" in close to 90 percent. At their second default, these percentages are 97 percent due to legal bankruptcy, and 6 percent to "distraint, no assets," in the cases of terminal defaults. Among the firms that experience a second but non-terminal default, 62 percent has the status of the "distraint, no assets."

A.2 Balance sheet data

By means of a two-step procedure, the six financial ratios were selected as: earnings before interest, depreciation, taxes and amortization (EBITDA) over total assets (TA) (earnings ratio); interest payments (IP) over the sum of interest payments and earnings before interest, depreciation, taxes and amortization (interest coverage ratio); total liabilities (TL) over total assets (leverage ratio); total liabilities over total sales (TS) (debt ratio); liquid assets (LA) in relation to total liabilities (quick ratio); and inventories (I) over total sales (inventory turnover ratio).[A.3]

[A.2] Firms that are declared bankrupt at some point do not disappear from the databases that UC maintain. Firm identifiers (*organisationsnummer*) are unique and are never re-cycled by Swedish tax authorities.

[A.3] It should be noted that the log-level of debt, in addition to the leverage ratio ($\mathrm{TL}_{i,t}/\mathrm{TA}_{i,t}$) for firm i in period t, contains predictive power for default. We therefore decided to include $\mathrm{TL}_{i,t}$ as a separate variable, but scaled it with average total sales in period t to obtain a stationary ratio. Thus, the debt-to-sales ratio is defined as $\log(\mathrm{TL}_{i,t}/\mathrm{TS}_t)$,where TS_t denotes average total sales in period t.

Table A.1 in Appendix A.3 provides an account of the variables considered by other well-known studies in the literature.

First, the univariate relationship between a ratio and default risk was investigated. By visual inspection, ratios that are largely uncorrelated with default risk were eliminated from the set of candidate explanatory variables. Figure A.1 illustrates this for the six selected ratios by comparing default rates (jagged line) and the cumulative distributions of the variables (smooth line) for all observations in the period $1990Q1-2009Q2$. The default rate for a given observation of a ratio is calculated as an average over the interval of +/- 5000 adjacent observations in the empirical distribution of the ratio at hand. The cumulative distribution at any point X_0 on the X-axis gives the share of defaulted firms for which the financial ratio is smaller than X_0. Given the density of the observations, there is a positive relationship between default and the leverage, interest coverage and turnover ratios, while there is a negative relationship for both the earnings and the liquidity ratios. Moreover, Figure A.1 suggests that the relationships between default and the earnings ratio, total liability over total sales ratio and interest costs over the sum of interest costs and earnings are all non-linear. For instance, for the interest coverage variable, this relationship is quite intuitive. The ratio turns highly negative if earnings are negative and slightly larger than interest payments in absolute value, which is associated with an increased risk of default. On the other tack, large interest payments and low earnings will also make the ratio large, positively, which is likewise associated with an increased default risk. Similar reasoning can be be applied to the other ratios. In the second step in the selection procedure, variables that did not enter significantly were subsequently dropped one by one to get the final set of variables. For instance, standard variables like size (proxied by total sales) and age (proxied by the number of periods in the panel) were dropped in this second step as they were found to be insignificant in the full model.

Regarding the definition of the dummy variable for firms that have not submitted a financial statement, TTLFS, there are three points worth noting. First, this information is assumed to be available with a 6-quarter time lag, since financial statements for year τ are typically available in the third quarter of year $\tau + 1$. By letting the dummy variable equal unity with a 6-quarter time-lag, we account for the real-time delay. Second, given the way we define the population of existing firms, recently registered firms entering the panel would automatically be assigned TTLFS = 1 in the third quarter of their existence, since they have not, of course, issued any financial statement prior to entering. For these new firms, TTLFS has been set to 0 and the

accounting data variables have been taken from their first yearly balance sheet and income statement. Third, for defaulting firms that are in the panel but have on no occasion submitted an annual report, we also set TTLFS equal to 0. This is the case for about 20 percent of the 161,550 defaulting firms in the panel. So, although TTLFS turns out to be very important in the default-risk models, this feature is down-played rather than exaggerated in the construction of the variable.

A.3 Descriptive statistics for winsorized data

In Table A.2, we report the means and standard deviations for a set of accounting ratios, payment remarks, and a variable that measures the average elapsed time since the latest filing of a financial statement for the final data set that is used in the estimations in Section 3.

The table distinguishes between defaulted and non-defaulted firms, at the aggregate as well as the industry level, for the in-sample period $1990Q1 - 1999Q4$, that is, the sub-sample period for which we will specify and estimate all subsequent models. For this period, we have a total of 8,106,176 observations of which 105,605 are defaults. Analyses of industry effects will be conducted at the one-digit level to ensure sufficiently many default observations in each industry in both the cross-sectional and the time series dimensions. The ten industries are; agriculture, manufacturing, construction, retail, hotel and restaurant, transportation, banking, finance and insurance, real-estate, consulting and rental, and finally a residual industry labeled "not classified".

Because of the varying availability of data, the statistics in Table A.2 are calculated based on slightly different numbers of observations for the variables in a given industry. Dealing with microdata sets of this size invariably involves dealing with outliers. These observations would distort the estimation results if they were to be included in the logit model and therefore, we have winsorized the top and bottom 1 percent observations for the accounting variables in each industry.[A.4] Given the large number of observations in our data set, this approach is practically more or less equivalent to simply deleting 1 percent of the observations that have accounting data that fall outside a certain region. Note that we choose to winsorize the observations in each industry separately, rather than at the aggregate level, thereby implicitly allowing for dispersion and different means in different industries. Table A.2 shows the descriptive statistics for the

[A.4] Winsorization is quite common in the literature using financial ratios to avoid outliers that are created by near-zero denominators. Shumway (2001) winsorizes the top and bottom 1 percent of all observations. It should be emphasized that the results are robust to varying the winsorization rate between 0.5 and 2 percent.

winsorized microdata set.[A.5]

A.4 Macro data

The aggregate time series are depicted in Figure A.2. They are: the output gap (i.e., the deviation of GDP from its trend value), the yearly inflation rate (measured as the fourth difference of the GDP deflator), the repo interest rate (a short-term nominal interest rate, set by the Riksbank), and the real exchange rate.

The repo rate was extremely high in the third quarter of 1992 due to the Riksbank having raised the so-called marginal interest rate to 500 percent, unexpectedly and temporarily, in an attempt to defend the fixed exchange rate. If the repo rate is not adjusted for this exceptional event, the estimation procedure would lead to underestimation of the importance of financial costs for default behavior. We therefore adjust the repo rate series in the third quarter of 1992 by means of a simple of a simple regression $R_t = b_1 + b_2 D923 + b_3 R_{t-1} + \varepsilon_t$. The estimated dummy coefficient $\hat{b}_2$ equals 28.2, and we therefore adjusted the repo rate $1992Q3$ to equal 9.8 percent instead of 38 percent.

The output gap series is computed by HP-filtering GDP, where the smoothing coefficient λ is set to the standard value of $1,600$. The real exchange rate is measured as the nominal TCW-weighted (TCW = trade competitive weights) exchange rate times the TCW-weighted foreign price level (CPI deflators) divided by the domestic CPI deflator. Note that a larger value for the real exchange rate implies a depreciation; hence a negative estimated coefficient for this variable implies that a depreciation on average reduces the risk of default at a given point in time. During the sample period, the real exchange rate is characterized by an upward trend (i.e.,a tendency of gradual depreciation) and it is therefore detrended with the HP-filter as well to achieve stationarity. Appendix B.3 verifies that the results are robust with respect to our detrending procedure of the real exchange rate.

In the output-gap series in Figure A.2 there is clear evidence of the deep recession in the beginning of the 1990s with a negative output gap of more than 4 percent in 1993. The economic rebound of 1994-1995 is also evident, as is the IT-boom bust cycle in the late 1990. Finally,

[A.5] Comparison of the descriptive statistics for unwinsorized data makes it clear that defaulted firms are disproportionally more affected when winsorizing all observations jointly. Since the PAYREMARK, TAXARREARS, PAYDIV and TTLFS are dummy variables that are unaffected by choice of winsorization procedure, a joint one could lead to underestimation of the importance of the accounting data variables in the default risk model relative to these dummy variables. To check the robustness of our chosen approach, we used an alternative approach where we truncated the healthy and defaulted firms separately. As expected, the estimation results of the default-risk model with this alternative winsorization suggest a somewhat larger role for the accounting ratios, but the overall picture remains the same.

Figure A.2 documents the exceptionally sharp downturn in the economy during late 2008 and the beginning of 2009.

Table A.1: Firm-specific explanatory variables used in papers focusing on the development of models of default risk

						RHS variables																	
Paper	Year	Sample	Firms	Model	LHS variable	Liquidity				Profitability and efficiency				Solvency and leverage			Size	Other			
Altman	1968	66	Listed	Discriminant	Bankruptcy	(CA-CL)/TA				RE/TA,	EBIT/TA,		TS/TA	MVE/TC							
Frydman et al.	1985	200	Listed	Discriminant	Bankruptcy	LA/TA,	CA/TA,	LA/CL,	CA/CL	CF/TD,	EBIT/TA,	NI / TA		MVE/TC,	Ln(IC)		Ln(TA)				
Shumway	2001	39,745	Listed	Panel Logit	Bankruptcy	(CA-CL)/TA			CA/CL	RE/TA,	EBIT/TA,	NI / TA,	TS/TA	MVE/TL,	TL/TA		Ln(MVE/market)	r_E-r_M	σ_E-σ_M		
Pesaran et al.	2006	[2,219]	Listed	Merton	Default					r_E				MVE		TE/TA		σ_E	CR		
Duffie et al.	2007	392,404	Listed	Hazard	Default					r_E								DtD			
Bonfim	2008	113,119	Mixed	Probit/Hazard	Loan default			LA/CL				NI/TA	TS/TA		TL/TA,	TE/TA		IR	TSG		
Bharath, Shumway	2008	1.016m	Listed	Merton	Distance to default							NI/TA		MVE	TD			σ_E			
""	"			Hazard	Time to default							NI/TA		MVE	TD			PD			

Note Table 2: Lists of variables reflect the main model presented in each paper. In cases where no preferred model" was presented, the list reflects variables with significant variables in any model.The number of observations in a paper can vary dependent on model specification. CA = Current Assets, CL = Current Liabilities, TA = Total Assets, RE = Retained Earnings, EBIT = Earnings Before Interest and Taxes, TS = Total Sales, MVE = Market Value of Equity, TC = Total Credit, LA = Liquid Assets, CF = Cash Flow, TD = Total Debt, NI = Net Income, IC = Interest Coverage, r_E = Return on Equity, r_M = Market Return on Equity, σ_E = Volatility of Stock Returns, σ_M = Volatility of Market Stock Returns, TE = Total Equity, CR = credit Rating, DtD = Distance to Default, IR = Investment Rate, TSG = Total Sales Growth, PD = Probability of default. The term quick assets has been named liquid assets in this table because no consensus on the exact difference between liquid and quick assets. Some sources include inventories in liquid assets, but exclude them from quick assets.

Table A.2: Descriptive statistics for truncated firm-specific micro data 1990*Q*1-1999*Q*4

	Agriculture	Manufacturing	Construction	Retail	Hotel & Resturant	Transport	Bank, Finance & Insurance	Real- Estate	Consulting & Rental	Not Classified	Total
Defaulted											
Number Obs	1643	13052	11843	32350	5457	5147	811	6470	14619	14213	105605
EBITDA/TA	-0.04 (0.35)	-0.01 (0.29)	0.01 (0.29)	-0.06 (0.35)	-0.12 (0.55)	0..01 (0.33)	-0.09 (0.66)	0.01 (0.27)	-0.03 (0.45)	-0.07 (0.45)	-0.04 (0.38)
TL/TA	1.09 (0.55)	0.97 (0.39)	0.96 (0.40)	1.04 (0.56)	1.20 (0.82)	1.01 (0.44)	1.12 (1.25)	1.02 (0.55)	0.94 (0.59)	0.99 (0.67)	1.01 (0.57)
LA/TL	0.14 (0.73)	0.13 (0.57)	0.18 (0.57)	0.17 (0.70)	0.20 (0.64)	0.19 (0.74)	0.43 (1.86)	0.18 (0.85)	0.34 (1.22)	0.36 (1.41)	0.22 (0.91)
I/TS	0.38 (0.84)	0.20 (0.31)	0.15 (0.42)	0.27 (0.45)	0.04 (0.06)	0.01 (0.04)	0.20 (1.33)	0.35 (1.54)	0.08 (0.25)	0.16 (0.48)	0.19 (0.56)
TL/TS	-1.12 (1.71)	-2.64 (1.78)	-1.64 (1.78)	-2.46 (1.69)	-1.72 (1.61)	-2.34 (1.70)	-2.56 (2.56)	-0.62 (2.19)	-1.69 (1.88)	-1.54 (2.18)	-1.98 (1.93)
IP/(IP+EBITDA)	0.25 (0.96)	0.24 (1.02)	0.18 (0.87)	0.23 (1.17)	0.19 (1.00)	0.23 (0.76)	0.23 (1.05)	0.41 (0.83)	0.15 (0.86)	0.20 (0.92)	0.22 (1.00)
PAYDIV	0.00 (0.03)	0.00 (0.05)	0.00 (0.05)	0.00 (0.04)	0.00 (0.04)	0.00 (0.03)	0.00 (0.04)	0.00 (0.04)	0.00 (0.05)	0.00 (0.03)	0.00 (0.04)
TTLFS	0.41 (0.49)	0.30 (0.46)	0.37 (0.48)	0.36 (0.48)	0.35 (0.48)	0.40 (0.49)	0.51 (0.50)	0.44 (0.50)	0.44 (0.50)	0.50 (0.50)	0.39 (0.49)
REMARK1	0.15 (0.36)	0.12 (0.32)	0.15 (0.36)	0.14 (0.35)	0.16 (0.37)	0.16 (0.37)	0.21 (0.41)	0.14 (0.35)	0.17 (0.37)	0.17 (0.37)	0.15 (0.36)
REMARK2	0.47 (0.50)	0.36 (0.48)	0.46 (0.50)	0.39 (0.49)	0.46 (0.50)	0.49 (0.50)	0.42 (0.49)	0.34 (0.47)	0.44 (0.50)	0.38 (0.49)	0.41 (0.49)
Non-defaulted											
Number Obs	259298	1229283	902494	2227966	258191	497011	120184	438139	1597081	470924	8000571
EBITDA/TA	0.14 (0.25)	0.12 (0.21)	0.12 (0.21)	0.08 (0.24)	0.08 (0.35)	0.17 (0.23)	0.08 (0.42)	0.09 (0.20)	0.14 (0.28)	0.13 (0.34)	0.11 (0.25)
TL/TA	0.70 (0.34)	0.69 (0.30)	0.71 (0.29)	0.74 (0.36)	0.87 (0.50)	0.72 (0.30)	0.68 (0.70)	0.80 (0.39)	0.63 (0.36)	0.65 (0.43)	0.71 (0.37)
LA/TL	0.48 (1.12)	0.43 (1.01)	0.44 (0.79)	0.43 (1.03)	0.38 (0.77)	0.46 (0.97)	1.42 (4.67)	0.42 (1.31)	0.90 (1.82)	0.99 (2.03)	0.58 (1.42)
I/TS	0.24 (0.57)	0.14 (0.24)	0.11 (0.31)	0.19 (0.34)	0.03 (0.06)	0.01 (0.05)	0.36 (1.80)	0.19 (1.11)	0.04 (0.18)	0.07 (0.34)	0.12 (0.44)
TL/TS	0.66 (1.30)	0.50 (1.85)	0.58 (1.81)	0.33 (0.90)	0.51 (1.25)	0.48 (1.74)	4.33 (26.14)	5.09 (20.15)	0.64 (2.26)	0.50 (1.58)	0.80 (6.03)
IP/(IP+EBITDA)	0.15 (0.71)	0.15 (0.72)	0.12 (0.72)	0.18 (0.87)	0.17 (0.79)	0.14 (0.54)	0.18 (0.94)	0.29 (0.68)	0.10 (0.75)	0.11 (0.75)	0.15 (0.77)
PAYDIV	0.14 (0.35)	0.15 (0.36)	0.13 (0.34)	0.13 (0.33)	0.06 (0.24)	0.13 (0.34)	0.13 (0.33)	0.08 (0.28)	0.16 (0.36)	0.12 (0.32)	0.13 (0.34)
TTLFS	0.00 (0.03)	0.00 (0.03)	0.00 (0.03)	0.00 (0.03)	0.00 (0.03)	0.00 (0.03)	0.00 (0.04)	0.00 (0.03)	0.00 (0.03)	0.00 (0.03)	0.00 (0.03)
REMARK1	0.00 (0.05)	0.00 (0.06)	0.00 (0.06)	0.00 (0.06)	0.01 (0.08)	0.00 (0.05)	0.00 (0.05)	0.00 (0.07)	0.00 (0.05)	0.00 (0.05)	0.00 (0.06)
REMARK2	0.03 (0.16)	0.03 (0.18)	0.04 (0.20)	0.03 (0.18)	0.06 (0.24)	0.04 (0.18)	0.02 (0.14)	0.03 (0.16)	0.03 (0.17)	0.03 (0.16)	0.03 (0.18)

Notes: The definition of variables are: EBITDA = earnings before taxes, interest payments and depreciations; TA = total assets; TL = total liabilities; LA = liquid assets; I = inventories; TS = total sales; IP = sum of net interest payments on debt and extra-ordinary net income; PAYDIV = a dummy variable equal 1 if the firm has paid out dividends during the accounting period and 0 otherwise; TTLFS = a dummy variable equal to 1 if the firm has not submitted an annual report in the previous year, and 0 otherwise; REMARK1 = a dummy variable taking the value of 1 if the firm has a payment remark due to one or more of the following events in the preceding four quarters; (i) a "non-performing loan" at a bank, or (ii) a bankruptcy petition, or (iii) issuance of a court order to pay a debt, or (iv) seizure of property; REMARK2 = a dummy variable taking the value of 1 if the firm is in various tax arrears.

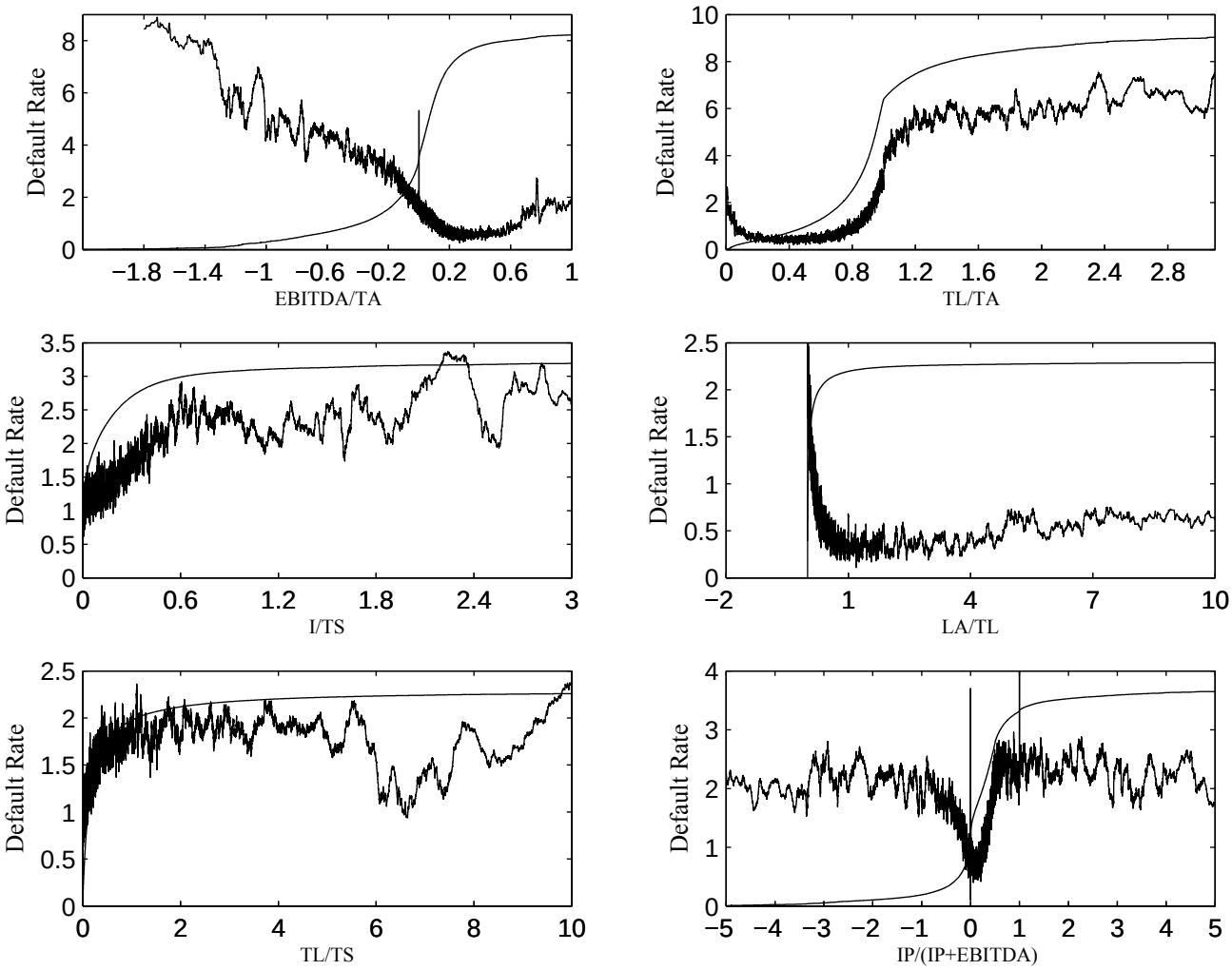

Figure A.1: Default rates and the cumulative distribution functions for the accounting data
The smooth lines are cumulative distributions of the default rate; the vertical axes on the right-hand-side have a minimum of zero and a maximum of one. Default frequencies at any point x_0 on the x-axis are computed as the average over 5,000 observations with values for the financial ratio smaller than x_0 and 5,0000 observations greater than x_0.

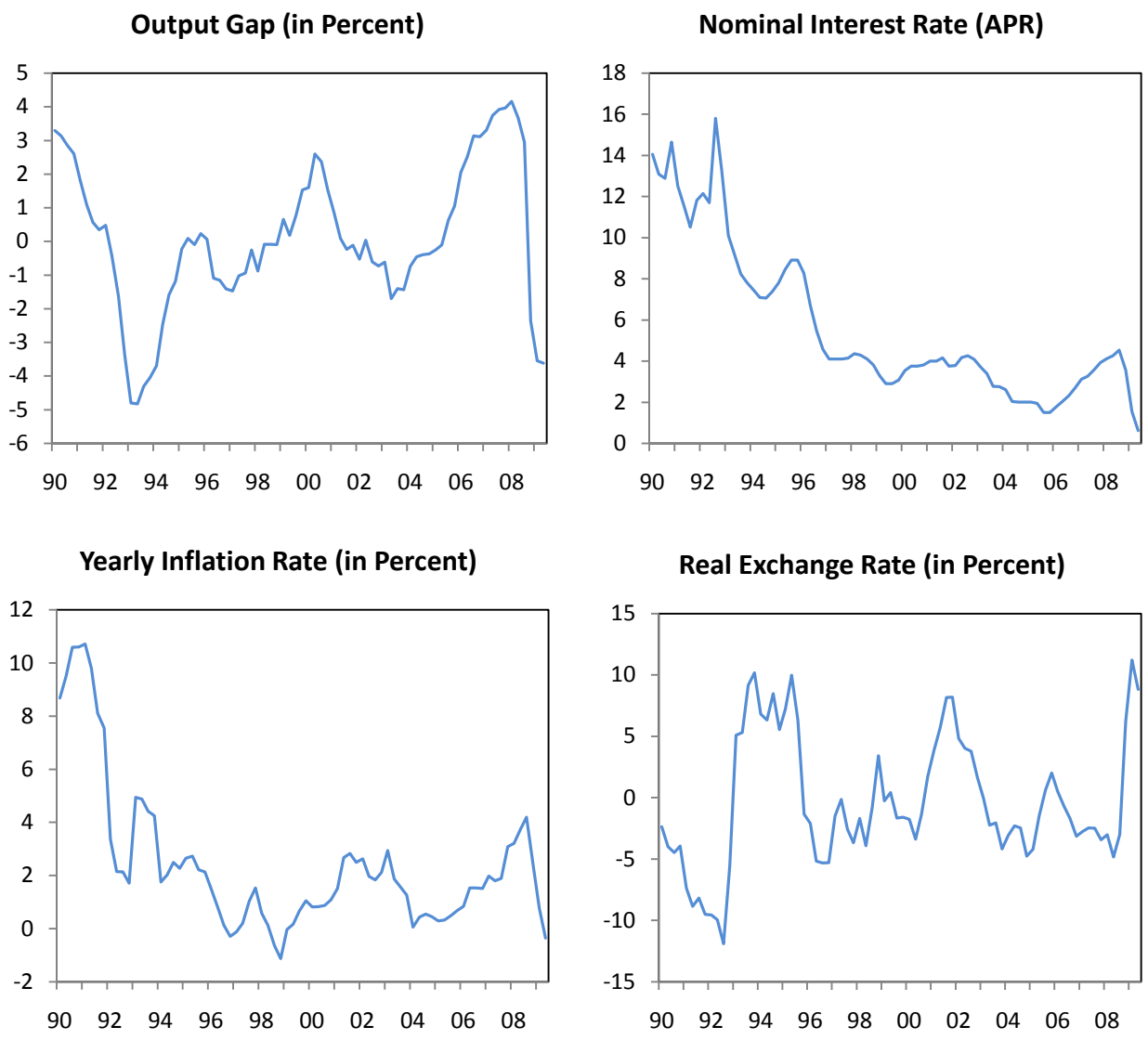

Figure A.2: Swedish Macro data 1990Q1-2009Q2 used in the estimated panel logit models

Appendix B Robustness analysis

The purpose of this appendix is to demonstrate that the results for the estimated models in Tables 1 and 2 are robust with respect to a number of perturbations. To keep the analysis tractable, we will restrict the analysis to the economy-wide models.

B.1 Role of remark data

Panel A in Table B.1 reports estimation results for the economy-wide models in Tables 1 and 2 when the PAYREMARK and TAXARREARS dummies have been dropped. For the sake of comparison, we also report the estimation results when the remarks variables are included in Panel A. As can be seen from comparing the results, the estimated coefficients for the accounting variables are similar irrespective of whether the remarks variables are included in the model or not. The coefficients for TL/TA and LA/TL increase when the remark variables are excluded, but the coefficient for EBITDA/TA is reduced. The remaining coefficients are roughly unaffected. Not in any case does exclusion of the Remark variables change the sign of the coefficients for the financial ratios. Thus, our estimation results for the financial ratios are not crucially affected by the inclusion of the remark variables. Nor do they diverge from the previous literature. However, it is clear that omitting the remark variables reduces the pseudo-R^2 measures, and thus reduces the ability of the model to rank the relative riskiness of firms. Turning to models where we include the macroeconomic variables in the regressions (c.f., Table 2), we again see that the coefficients are quite similar. An exception is the output gap coefficient, which turns out to be somewhat smaller in the model without remark variables. However, the coefficients for the output gap and the nominal interest rate are still of key importance, and thus the overall roles of macroeconomic variables are not affected by the presence of remark data. Once again, we find that extending the set of firm-specific factors with the remark variables allows us to increase the pseudo-R^2 measures substantially.

We conclude from this analysis that inclusion of remark data is not of key importance for the estimated impact of macroeconomic factors. Accordingly, our findings regarding the impact of macroeconomic factors would therefore be expected to hold in other countries, where payment remark data is not available. Nonetheless, the models' exceptional risk-ranking performance (as documented in Section 4.2) is clearly partly driven by the possibility to include remark data; without these data the out-of-sample risk-ranking performance would be worse than the in-

sample results reported by Shumway (2001) for publicly listed firms. But, given that the firms in our data set (i.e., the entire population of Swedish firms) are very heterogeneous, it is still a remarkable that the risk-ranking performance of the models is quite acceptable even without the remark data.

B.2 Imputation of missing financial ratios

Panel B in Table B.1 reports estimation results for the economy-wide model in Tables 1 and 2 for regressions when only firms that have submitted complete financial statements for all sample periods are selected. In this case, obviously no imputation of missing financial ratios is carried out and the interaction dummy TTLFS is dropped since it will equal 0 for all firms. In addition, we drop all firms that have been active for a too short time period to meet the financial statement requirement. Comparing the results in Panel B with the benchmark results in Panel A, we see that the coefficients are little affected by our imputation procedure. The absolute values for the coefficients for EBITDA/TA and for LA/TL increase somewhat, while the remaining financial ratios are more or less the same. The coefficients for the remark variables, and for the DIVIDEND-variable are roughly the same, as are the coefficients for the macroeconomic variables. One interesting conclusion from this robustness analysis is that TTLFS is even more important than the remark data for the fit of the model at the firm level (pseudo-R^2). In the full benchmark model (III) in Panel A, the pseudo-R^2 is 0.34. In the model without remark data the pseudo-R^2 falls to 0.24. Excluding TTLFS (i.e., information on whether the firm issued a financial report in due time or not) leads to an even larger fall in pseudo-R^2 to about 0.17, suggesting that this indicator variable is the single most important predictor of default. This result has a lot of intuitive appeal, one would think that failure to complile a financial report should be a very serious signal that a firm is in the sort of trouble that could lead to a permanent exit.

B.3 Data frequency, sample period, and the real exchange rate

First we consider the effects of excluding the real exchange rate in the Table 2 model. In addition, we report results when the real exchange rate is calculated as the percentage deviation around a constant mean ($q_t = (Q_t - \bar{Q})/\bar{Q}$, where $\bar{Q} = \frac{1}{T}\sum_{t=1}^{T} Q_t$ for the period $1990Q1 - 2009Q2$) instead of being HP-filtered. The results of these experiments are reported in Panel C in Table B.1. By and large, our results are not much affected by the choice of procedure for detrending

the real exchange rate. Differences in estimated coefficients for the financial ratios occur at the fourth decimal and are really miniscule. They are somewhat larger, though still small, for the set of indicator variables. For the other three macro variables we find that using the unfiltered real exchange rate reduces their coefficients marginally, while the real exchange rate coefficient is slightly increased.We conclude that using the filtered or unfiltered real exchange rate is of little consequence for the results in this paper. By a balanced regression argument we prefer the filtered real exchange rate as the benchmark, and we note in Section 2 that the results are robust w.r.t. to the detrending procedure.

In Panel D, we report results when we have estimated the model in Table 2 on an annual frequency instead of the quarterly frequency used in the paper. Again, we conclude that, on the whole, it is immaterial for the estimated parameters whichever choice of frequency is made. Noticeable exceptions are the coefficients for the debt ratio, TL/TA, and the output-gap, which both turn out stronger, while the coefficient for the real exchange rate is reduced, and that of inflatation (imprecisely estimated) switches sign. Therefore, a quarterly model seems more appealling since it allows for more detailed forecasting and more interesting interpretations. It is nevertheless reassuring that our disaggregation procedure does not seem to introduce unwarranted biases.

Finally, to examine the stability of the coefficients for the Economy Wide models of Table 1 and 2, we have re-estimated them using the full sample period, $1990Q1 - 2009Q2$, and present the results in Panel E in Table B.1. Comparing with the coefficients Panel E with those in Tables 1 and 2, we find only minor differences. The estimates for both the firm-specific and aggregate regressors are remarkably stable, consistent with the favorable out-of-sample results in Section 4 in both the cross-section and time series dimension.

B.4 Marginal effects

The explanatory variables in this paper have not been re-scaled to have the same mean, and therefore one cannot judge the importance of a particular variable from the size of its coefficient. The discussion about relative importance of explanatory variables in this paper is based on the relative sizes of the estimated t-statistics, since coefficient size is not sufficient for such inference. Alternatively, one can calculate the marginal contribution, or effect, from a variable at the mean, or the median, of the variable. Table B.2 report on such marginal effects and the calculations yield similar rankings of importance as the standard t-statistics, e.g., the output-gap and the

nominal interest rate are the more influential macroeconomic variables.

Table B.1: Sensitivity analysis for the economy-wide default risk models in Tables 1 and 2 1990Q1-1999Q4

	Panel A: Robustness w.r.t. Remarks				**Panel B:** Robustness w.r.t. TTLFS[a]				**Panel C:** Robustness w.r.t. real ex. rate			**Panel D:** Robustness w.r.t. freq		**Panel E:** Robustness w.r.t. Sample	
	Original sample (as in Tables 1 and 2)				*Only include firms who have always submitted all FS data*				*Use alternative RER and omit RER*			*Quarterly vs. Annual Data*			
									Benchmark	Alt. QD	Omit QD	Quarterly	Annual	Full Sample (90Q1 - 09Q2)	
	I	II	III	IV	I	II	III	IV	I	II	III	I	II	I	II
Micro Variables															
EBITDA/TA	**-0,8371**	**-0,7594**	**-0,8221**	**-0,7477**	**-0,9929**	**-0,9082**	**-0,9747**	**-0,8919**	**-0,8221**	**-0,8222**	**-0,8259**	**-0,8221**	**-0,9411**	**-0,7139**	**-0,7547**
	0,0106	*0,0098*	*0,0107*	*0,0098*	*0,0125*	*0,0117*	*0,0125*	*0,0118*	*0,0107*	*0,0107*	*0,0107*	*0,0107*	*0,012*	*0,0080*	*0,0080*
TL/TA	**0,3948**	**0,6982**	**0,3832**	**0,6976**	**0,3968**	**0,6926**	**0,3857**	**0,6977**	**0,3832**	**0,3834**	**0,3752**	**0,3832**	**0,497**	**0,3200**	**0,2982**
	0,0071	*0,0063*	*0,0072*	*0,0064*	*0,0083*	*0,0074*	*0,0084*	*0,0075*	*0,0072*	*0,0072*	*0,0072*	*0,0072*	*0,008*	*0,0053*	*0,0054*
LA/TL	**-0,2067**	**-0,3159**	**-0,2045**	**-0,3142**	**-0,4073**	**-0,6328**	**-0,3915**	**-0,6206**	**-0,2045**	**-0,2044**	**-0,2060**	**-0,2045**	**-0,194**	**-0,1517**	**-0,1425**
	0,0061	*0,0067*	*0,0061*	*0,0067*	*0,0121*	*0,0135*	*0,0121*	*0,0135*	*0,0061*	*0,0061*	*0,0061*	*0,0061*	*0,006*	*0,0037*	*0,0037*
I/TS	**0,0573**	**0,0682**	**0,0444**	**0,0607**	**0,0622**	**0,0718**	**0,0490**	**0,0633**	**0,0444**	**0,0442**	**0,0427**	**0,0444**	**0,047**	**0,0875**	**0,0655**
	0,0052	*0,0046*	*0,0053*	*0,0046*	*0,0059*	*0,0052*	*0,0060*	*0,0052*	*0,0053*	*0,0053*	*0,0053*	*0,0053*	*0,006*	*0,0043*	*0,0045*
TL/TS	**0,1078**	**0,0892**	**0,1009**	**0,0852**	**0,0913**	**0,0740**	**0,0880**	**0,0724**	**0,1009**	**0,1009**	**0,1015**	**0,1009**	**0,104**	**0,0823**	**0,0726**
	0,0020	*0,0018*	*0,0020*	*0,0018*	*0,0024*	*0,0023*	*0,0025*	*0,0024*	*0,0020*	*0,0020*	*0,0020*	*0,0020*	*0,002*	*0,0016*	*0,0016*
IP/(IP+EBITDA)	**0,0655**	**0,0737**	**0,0558**	**0,0674**	**0,0595**	**0,0638**	**0,0500**	**0,0567**	**0,0558**	**0,0557**	**0,0553**	**0,0558**	**0,067**	**0,0651**	**0,0515**
	0,0040	*0,0039*	*0,0039*	*0,0038*	*0,0046*	*0,0046*	*0,0044*	*0,0045*	*0,0039*	*0,0039*	*0,0039*	*0,0039*	*0,005*	*0,0036*	*0,0035*
PAYREMARK	**1,7256**		**1,8497**		**1,6866**		**1,8043**		**1,8497**	**1,8527**	**1,8467**	**1,8497**	**1,834**	**1,8976**	**1,9523**
	0,0145		*0,0147*		*0,0165*		*0,0168*		*0,0147*	*0,0147*	*0,0147*	*0,0147*	*0,017*	*0,0102*	*0,0102*
TAXARREARS	**2,5654**		**2,6839**		**2,3323**		**2,4713**		**2,6839**	**2,6857**	**2,6593**	**2,6839**	**2,282**	**2,6088**	**2,7241**
	0,0092		*0,0094*		*0,0108*		*0,0111*		*0,0094*	*0,0094*	*0,0094*	*0,0094*	*0,010*	*0,0070*	*0,0071*
Dividend	**-3,1728**	**-3,5374**	**-3,0066**	**-3,4573**	**-3,1324**	**-3,4098**	**-2,9531**	**-3,3082**	**-3,0066**	**-3,0046**	**-3,0436**	**-3,0066**	**-3,221**	**-3,0896**	**-2,9452**
	0,0708	*0,0706*	*0,0709*	*0,0706*	*0,0736*	*0,0736*	*0,0736*	*0,0736*	*0,0709*	*0,0709*	*0,0708*	*0,0709*	*0,068*	*0,0456*	*0,0456*
TTLFS	**3,6937**	**3,7990**	**3,6076**	**3,7532**					**3,6076**	**3,6133**	**3,6164**	**3,6076**	**3,082**	**3,9342**	**3,7527**
	0,0084	*0,0074*	*0,0085*	*0,0075*					*0,0085*	*0,0085*	*0,0085*	*0,0085*	*0,010*	*0,0073*	*0,0074*
Aggr. Variables															
Output gap			**-0,1327**	**-0,0782**			**-0,1460**	**-0,1061**	**-0,1327**	**-0,1205**	**-0,1076**	**-0,1327**	**-0,210**		**-0,1119**
			0,0026	*0,0024*			*0,0030*	*0,0029*	*0,0026*	*0,0025*	*0,0025*	*0,0026*	*0,004*		*0,0018*
Nominal interest rate			**0,0731**	**0,0501**			**0,0594**	**0,0351**	**0,0731**	**0,0522**	**0,0971**	**0,0731**	**0,065**		**0,0703**
			0,0016	*0,0015*			*0,0018*	*0,0018*	*0,0016*	*0,0020*	*0,0014*	*0,0016*	*0,002*		*0,0012*
GDP inflation			**0,0116**	**-0,0173**			**0,0363**	**0,0149**	**0,0116**	**0,0075**	**-0,0231**	**0,0116**	**-0,061**		**0,0158**
			0,0023	*0,0022*			*0,0026*	*0,0026*	*0,0023*	*0,0022*	*0,0020*	*0,0023*	*0,004*		*0,0020*
Real exchange rate			**-0,0258**	**-0,0071**			**-0,0341**	**-0,0198**	**-0,0258**	**-0,0272**		**-0,0258**	**-0,014**		**-0,0214**
			0,0008	*0,0003*			*0,0009*	*0,0009*	*0,0008*	*0,0008*		*0,0008*	*0,001*		*0,0007*
Mean log-likelihood	-0,046	-0,053	-0,046	-0,053	-0,039	-0,043	-0,039	-0,043	-0,046	-0,046	-0,046	-0,046	-0,125	-0,034	-0,034
Pseudo R2	0,33	0,23	0,34	0,24	0,16	0,08	0,17	0,08	0,34	0,34	0,34	0,34	0,35	0,36	0,37
Number of obs	8,106,176	8,106,176	8,106,176	8,106,176	7,949,015	7,949,015	7,949,015	7,949,015	8,106,176	8,106,176	8,106,176	8,106,176	2,207,382	16 928 521	16 928 521

Notes: Coefficient estimates in bold style, standard errors in italics. The variables have not been scaled, so the importance of a variable cannot be interpreted directly from the size of the parameter estimate. The pseudo R^2 values have been calculated according to McFadden (1974). [a] The number of observations in these estimations are reduced by 157,161 as we are only including firms for which financial statement reports are available. This means that we only include firms for which TTFLS = 0, and omit all defaulting firms that have never submitted financial accounting data. For reasons explained in more detail in footnote ?? in the paper, there are a large number of firms (27,492) that have never reported accounting data before they default, and these firms are all assigned TTLFS=0 in our analysis. Also these firms are excluded In the sub-sample considered here, so the number of defaults are only 64,189 in these estimations (as opposed to the 105,605 defaulting firms in Tables 1 and 2).

Table B.2: Marginal effects for the economy-wide default risk model

	Average of individual marginal effects				Marginal effects at the mean				Marginal effects at the median			
	I	II	III	IV	I	II	III	IV	I	II	III	IV
Micro Variables												
EBITDA/TA	-0,0083	-0,0084	-0,0081	-0,0083	-0,0030	-0,0037	-0,0020	-0,0025	-0,0037	-0,0063	-0,0023	-0,0040
TL/TA	0,0039	0,0077	0,0038	0,0077	0,0014	0,0034	0,0009	0,0024	0,0017	0,0058	0,0011	0,0038
LA/TL	-0,0020	-0,0035	-0,0020	-0,0035	-0,0007	-0,0015	-0,0005	-0,0011	-0,0009	-0,0026	-0,0006	-0,0017
I/TS	0,0006	0,0008	0,0004	0,0007	0,0002	0,0003	0,0001	0,0002	0,0003	0,0006	0,0001	0,0003
TL/TS	-2,2790	-2,2790	-2,2790	-2,2790	0,0004	0,0004	0,0003	0,0003	0,0005	0,0007	0,0003	0,0005
IP/(IP+EBITDA)	0,1507	0,1507	0,1507	0,1507	0,0002	0,0004	0,0001	0,0002	0,0003	0,0006	0,0002	0,0004
PAYREMARK	0,0171		0,0181		0,0062		0,0046		0,0076		0,0053	
TAXARREARS	0,0254		0,0263		0,0092		0,0067		0,0113		0,0076	
Dividend	-0,0314	-0,0392	-0,0295	-0,0382	-0,0113	-0,0170	-0,0075	-0,0117	-0,0139	-0,0293	-0,0085	-0,0187
TTLFS	0,0365	0,0421	0,0353	0,0415	0,0132	0,0183	0,0089	0,0128	0,0162	0,0314	0,0102	0,0203
Aggr. Variables												
Output gap			-0,0013	-0,0009			-0,0003	-0,0003			-0,0004	-0,0004
Nominal interest rate			0,0007	0,0006			0,0002	0,0002			0,0002	0,0003
GDP inflation			0,0001	-0,0002			0,0000	-0,0001			0,0000	-0,0001
Real exchange rate			-0,0003	-0,0001			-0,0001	0,0000			-0,0001	0,0000

Notes: Marginal effects for the explanatory variables in the economy-wide model. No standard errors are shown: statistical significance of the marginal effects is at the same level as for the parameter estimates in Table 3. Model III corresponds to the economy-wide model in Table 3, while the marginal effects in the columns marked I, II and IV refer to the models estimated for robustness purposes and displayed in Table B1.