Empirical Appendix

In this appendix, we first describe in detail the procedure for cleaning the data in the regressions. We then examine the comparability of the country samples. Finally, we present additional robustness of the main empirical regressions.

Data Cleaning Procedure

We now describe the detailed procedures in assembling the cross-country datasets analyzed in the empirical section. In particular, we present step-by-step data cleaning procedures, construction methods of all variables, and data sources for the country-level statistics.

Firm Data

We download the data from the AMADEUS database compiled by Bureau Van Dijk Electronic Publishing for years 1999–2005. We first delete all firms in the financial and government sectors which correspond to NACE codes 65, 66, 67, and 75. For each cross-section regression of year t, we then restrict the sample to firms that report positive assets and nonnegative liabilities in year t and nonnegative sales in year t and $t+1$. We next drop firms with either leverage or sales growth at the top one percentile of the leverage or sales growth distribution of each country. We finally drop the countries with less than one thousand observations after the above cleaning procedures.

We construct all firm-level variables in year t as follows. For each firm, we generate the *leverage* variable by taking the ratio of the firm's total debt to total assets in year t. We generate the *Size* variable by taking logs of the share of the firm asset in its country in year t. We generate the interaction variables by multiplying *Size* by the corresponding variables of interest such as private credit to GDP or credit bureau coverage. We generate the *growth* variable by computing the firm' net sales growth rate from year t to $t+1$. Given the country-level fixed effect in all the regressions, we need not adjust the sales growth with the country-level inflation rate and changes in the exchange rate. We construct dummy variables for age groups. Firms are classified into 7 age groups based on the firm age in terms of years: $[0,5)$, $[5,10)$, $[10,15)$, $[15,20)$, $[20,25)$, $[25,30)$, $[30,\infty)$. We construct the *entry* dummy to be one for firms with age less than or equal to two.

Country Data

The country-level statistics are obtained from various data sources. Private credit to GDP is from the *World Development Indicators* of the World Bank. Credit bureau coverage is from various issues of *Doing Business* of the World Bank. Overhead costs are taken from Beck et al. (2009). CPI inflation rates are from the *International Financial Statistics* of the International Monetary Fund.

Comparability of Country Samples

This section analyzes the coverage and comparability of the AMADEUS dataset across countries. Eurostat 2007 contains information on the distribution of the universe of firms in the business sector for most of the countries in our sample in 2004. They report the percentage of enterprises that have

1–49 employees, 50–250 employees, and over 250 employees. Hence, we compare the fraction of firms for each employment category in our dataset with that in the universe from the report.[1]

Unfortunately, the employment information is not reported for every firm in AMADEUS. On average, about 66% of firms in our cleaned sample report employment statistics in 2004. The lack of employment data can be a severe problem for some countries. For example, only 21% of firms in our cleaned sample of Italy report employment. Moreover, this lack of employment information is the most severe for small firms. Hence, we impute employment measures for firms that do not report employment in AMADEUS. To do this, we run regressions country by country of log(employment) on log(assets) and log(sales). The fit of these regressions is good, with R^2 above 0.6 for all countries.[2] We then impute employment for the firms that do not report it using the estimated coefficients, their assets and sales.

Table 1 reports the firm distribution in AMADEUS and in the universe for countries for which we have data. The table shows that in our sample, the majority of firms are small with less than 49 employees as in the data. The coverage of small firms in our sample, though large (on average about 80%), is low relative to that in the universe of firms (on average about 98%). In our sample, about 7% of firms have more than 250 employees, but in the universe, less than 1% of firms fall into this category. On the other hand, the coverage in AMADEUS is similar across countries, with 18 out of 27 countries having small firms larger than 80%.

Robustness of the Main Regressions

Tables 2 and 3 report the regression coefficients when the additional interactions are added to control for the alternative explanations with financial development measured as overhead costs and credit bureau coverage. We find in both the leverage and growth regressions that the coefficients of the interaction between size and financial development remain the same sign as those in the main regressions, except for one specification.[3] In the growth regressions, the coefficients of the interaction terms with the entry dummy are almost the same as those in the main results. The same is true for the leverage regressions, except for the overhead-costs case where the coefficient of the interaction between size and entry has the opposite sign.

Table 4 and 5 report the regression coefficients for the previous sample years with financial development measured as overhead costs and credit bureau coverage. Across all the sample years, the size-growth regression generates all the coefficients similar to our main results, in terms of both the magnitude and the statistical significance. Thus, regardless of the measure of financial development, our main results on the cross-country size-growth relation are robust over time. For the size-leverage regression, the coefficients of size and the interaction terms of size with financial development are similar to those in the main regression across all the years. The only less robust coefficient is the interaction terms with the entry due to low and spurious coverage of entrants in the early years.

---

[1]For this comparison, we include only firms in sectors that correspond to the business sectors in the European Commission Report.

[2]Introducing additional controls such as firm age and sector dummies changes the fit of the regressions only marginally.

[3]When all three interactions are included, the growth regression has the opposite sign for the coefficient of the interaction between firm size and credit bureau coverage.

Computation Appendix

In this appendix, we provide the computation algorithm. The state variables are capital K, debt repayment B_R, and productivity z. We first discretize the state variables. We then start with an initial debt schedule in which all contracts have a risk-free interest rate. We then solve the model using the value function iterations. To do so, we make an initial guess over repayment value $V^c(K, B_R, z)$ and default decision $d(K, B_R, z)$, which provides continuation value function $V(K, B_R, z)$. We then solve for the decision rules by searching over grid points. Consequently, we obtain an updated repayment value, an updated default function, and the associated decision rules. We continue with the above procedures until the repayment value function and the default decision function converge. Now we update the debt schedule with the default decision. We repeat these processes until the debt schedules converge.

Given the complexity of our computation, the stochastic shock process is assumed to be a two-point Markov process, and the permanent productivity process has five points. The two stochastic shock levels and the symmetric transition matrix between these two points are set to match a mean of one without loss of generality and the empirical estimates of the shock persistence and volatility. The five permanent productivity levels are set to match the size distribution in the data. We use 200 capital grids and 200 debt grids, which are large enough to provide a reliable solution. Doubling the numbers of the grid points produces almost identical results.

Sensitivity with Proportional Credit Costs

In this appendix we conduct an experiment in which the fixed credit cost is set to zero, and firms pay instead a financing cost, which is proportional to loan size. Countries with better financial development have smaller proportional financing costs for borrowers with saving rates unchaged. Worsening financial development in this case is equivalent to increasing the risk free interest rate for *loans*. To see this, the creditors' break-even condition is given by

$$(B' + \chi B')(1 + r) = \left(1 - \int d(K', B_R', z'))f(z', z)dz'\right) B_R' + \int d(K', B_R', z'))f(z', z)dz' R(K').$$

where χ is the proportional cost of financing. The effective interest rate schedule is given by

$$\frac{B_R'}{B'} = \frac{(1 + \chi)(1 + r)}{1 - \int d(K', B_R', z'))f(z', z)dz'} - \frac{\int d(K', B_R', z'))f(z', z)dz' R(K')}{(1 - \int d(K', B_R', z'))f(z', z)dz')B'}.$$

Thus, varying χ is equivalent to varying r in the model. We set the fixed credit cost to zero, and experiment with two lending rates: 6% and 4%. The rest of the parameters are the same as in the benchmark calibration. In particular, the savings rate is set at 4% in both experiments. See Table 6 for results. We find that as the interest rate lowers and financial development improves, the model implies that the growth difference between small and large firms reduces substantially from 54% to 8%, which is consistent with the empirical findings. However, the leverage difference decreases from 0% to -6%, which is inconsistent with our empirical findings. Thus, a proportional financing cost will not be able to reconcile the cross-country and cross-firm patterns of firm financing observed in the data. In contrast, the presence of fixed credit costs allows the model to match our empirical findings.

Table 1: Coverage and Comparability of Country Datasets

	AMADEUS Dataset			EC Universe		
	Small	Medium	Large	Small	Medium	Large
	1-49	50-250	> 250	1-49	50-250	> 250
Belgium	0.897	0.076	0.027		0.009	
Bosnia	0.844	0.130	0.025			
Bulgaria	0.898	0.078	0.024	0.982	0.016	0.002
Croatia	0.972	0.023	0.004			
Czech	0.854	0.114	0.031	0.991	0.008	0.001
Estonia	0.971	0.026	0.004	0.966	0.030	0.004
Finland	0.910	0.062	0.028	0.985	0.012	0.003
France	0.963	0.029	0.008	0.987	0.010	0.003
Germany	0.617	0.212	0.171	0.972	0.023	0.005
Greece	0.892	0.092	0.016			
Hungary	0.991	0.008	0.001			0.001
Iceland	0.989	0.010	0.001			
Ireland	0.523	0.227	0.249			
Italy	0.952	0.039	0.010	0.994	0.005	0.001
Latvia	0.690	0.242	0.068	0.970	0.027	0.003
Lithuania	0.495	0.396	0.108	0.952	0.043	0.005
Netherlands	0.329	0.354	0.317	0.981	0.016	0.003
Poland	0.353	0.453	0.194	0.989	0.009	0.002
Portugal	0.844	0.125	0.031			
Romania	0.973	0.022	0.005	0.971	0.023	0.006
Russia	0.806	0.138	0.056			
Serbia	0.863	0.096	0.042			
Slovakia	0.684	0.239	0.077		0.055	0.014
Spain	0.959	0.034	0.008	0.991	0.008	0.001
Sweden	0.940	0.046	0.014	0.990	0.008	0.002
Ukraine	0.298	0.446	0.256			
United Kingdom	0.738	0.163	0.098	0.978	0.018	0.004
AVERAGE	0.787	0.144	0.069	0.980	0.019	0.004

Table 2: Robustness with Additional Interactions: Overhead Costs

	Leverage		Sales Growth	
$FD \times Size$	0.054***	0.021***	−0.986***	−0.366***
	(0.0058)	(0.0062)	(0.0328)	(0.0355)
$Entry \times Size$	−0.000**	−0.000	0.064***	0.060***
	(0.0001)	(0.0001)	(0.0014)	(0.0014)
$Entry \times FD \times Size$	−0.004	−0.010***	−2.249***	−2.139***
	(0.0030)	(0.0030)	(0.0368)	(0.0362)
$GDPVOL \times Size$	−0.010***	−0.072***	−0.877***	0.271***
	(0.0038)	(0.0053)	(0.0223)	(0.0309)
$GDP \ per \ capita \times Size$		−0.002***		0.042***
		(0.0001)		(0.0009)
$Industry \times Size$	Yes	Yes	Yes	Yes
Adjusted R^2	0.28	0.28	0.07	0.07
Number of observations	2606324	2606324	2606324	2606324

Notes: *Size* is measured by total asset of a firm. *FD* denotes financial development, measured by overhead costs. *Entry* is a dummy variable, which equals one for new entrants defined as firms with age less than or equal to two years. *GDPVOL* is the standard deviation of GDP growth rates. All regressions have a fixed effect at the country×industry×age level. The standard errors reported in parentheses are robust to heteroskedasticity. *** denotes significant at 1%. ** denotes significant at 2%.

Table 3: Robustness with Additional Interactions: Credit Bureau Coverage

	Leverage		Sales Growth	
$FD \times Size$	-0.003^{***}	-0.002^{***}	0.025^{***}	-0.003^{***}
	(0.0003)	(0.0003)	(0.0008)	(0.0008)
$Entry \times Size$	-0.001^{***}	-0.001^{***}	-0.039^{***}	-0.038^{***}
	(0.0000)	(0.0000)	(0.0004)	(0.0004)
$Entry \times FD \times Size$	0.001^{***}	0.001^{***}	0.035^{***}	0.033^{***}
	(0.0002)	(0.0002)	(0.0008)	(0.0008)
$GDPVOL \times Size$	-0.015^{***}	-0.085^{***}	-1.038^{***}	0.470^{***}
	(0.0034)	(0.0050)	(0.0214)	(0.0297)
$GDP\ per\ capita \times Size$		-0.002^{***}		0.052^{***}
		(0.0001)		(0.0008)
$Industry \times Size$	Yes	Yes	Yes	Yes
Adjusted R^2	0.28	0.28	0.07	0.07
Number of observations	2621201	2621201	2621201	2621201

Notes: *Size* is measured by total asset of a firm. *FD* denotes financial development, measured by private credit bureau coverage. *Entry* is a dummy variable, which equals one for new entrants defined as firms with age less than or equal to two years. *GDPVOL* is the standard deviation of GDP growth rates. All regressions have a fixed effect at the country×industry×age level. The standard errors reported in parentheses are robust to heteroskedasticity. *** denotes significant at 1%.

Table 4: Robustness with Additional Years: Overhead Costs

	1999	2000	2001	2002	2003
			Leverage		
Size	0.014***	0.015***	0.015***	0.014***	0.014***
	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0003)
FD×Size	0.137***	0.129***	0.127***	0.144***	0.033***
	(0.0053)	(0.0060)	(0.0065)	(0.0064)	(0.0053)
Entry×Size	0.000	−0.000	−0.001***	−0.002***	0.001***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Entry×FD×Size	−0.022***	−0.010*	0.001	0.051***	−0.027***
	(0.0047)	(0.0052)	(0.0051)	(0.0045)	(0.0037)
Adjusted R^2	0.20	0.17	0.17	0.21	0.22
			Sales Growth		
Size	0.060***	0.082***	0.064***	0.045***	0.032***
	(0.0020)	(0.0024)	(0.0023)	(0.0020)	(0.0014)
FD×Size	−2.493***	−3.016***	−2.701***	−2.252***	−2.096***
	(0.0574)	(0.0632)	(0.0602)	(0.0590)	(0.0404)
Entry×Size	0.050***	0.052***	0.057***	0.049***	0.063***
	(0.0024)	(0.0025)	(0.0024)	(0.0019)	(0.0018)
Entry×FD×Size	−1.825***	−1.919***	−2.038***	−1.882***	−2.278***
	(0.0680)	(0.0690)	(0.0671)	(0.0529)	(0.0507)
Adjusted R^2	0.09	0.08	0.08	0.07	0.08
Observations	1269539	1396303	1582012	1842944	2108389

Note: *Size* is measured by total asset of a firm. *FD* is measured as overhead costs. *Entry* is a dummy variable, which equals one for new entrants defined as firms with age less than or equal to two years. All regressions have a fixed effect at the country×industry×age level. The standard errors reported in parentheses are robust to heteroskedasticity. *** denotes significant at 1%, ** denotes significant at 5%, and * denotes significant at 10%.

Table 5: Robustness with Additional Years: Credit Bureau Coverage

	1999	2000	2001	2002	2003
			Leverage		
Size	0.022***	0.023***	0.025***	0.022***	0.017***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0001)
$FD \times Size$	−0.007***	−0.011***	−0.014***	−0.003***	−0.006***
	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0003)
$Entry \times Size$	−0.001***	−0.001***	−0.001***	−0.000***	−0.001***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$Entry \times FD \times Size$	−0.001*	−0.001**	0.001**	−0.000	0.002***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0002)
Adjusted R^2	0.20	0.17	0.17	0.21	0.22
			Sales Growth		
Size	−0.081***	−0.087***	−0.091***	−0.080***	−0.087***
	(0.0017)	(0.0018)	(0.0015)	(0.0015)	(0.0013)
$FD \times Size$	0.074***	0.086***	0.085***	0.054***	0.056***
	(0.0018)	(0.0019)	(0.0016)	(0.0012)	(0.0010)
$Entry \times Size$	−0.030***	−0.033***	−0.031***	−0.034***	−0.037***
	(0.0007)	(0.0006)	(0.0006)	(0.0005)	(0.0005)
$Entry \times FD \times Size$	0.018***	0.024***	0.024***	0.028***	0.031***
	(0.0013)	(0.0013)	(0.0012)	(0.0010)	(0.0010)
Adjusted R^2	0.08	0.08	0.07	0.07	0.07
Observations	1282817	1409644	1595747	1857510	2124784

Note: *Size* is measured by total asset of a firm. *FD* is measured as credit bureau coverage. *Entry* is a dummy variable, which equals one for new entrants defined as firms with age less than or equal to two years. All regressions have a fixed effect at the country×industry×age level. The standard errors reported in parentheses are robust to heteroskedasticity. *** denotes significant at 1%, ** denotes significant at 5%, and * denotes significant at 10%.

Table 6: Sensitivity on Proportional Credit Costs

	Model with $r=6\%$		Model with $r=4\%$	
All firms	Growth	Leverage	Growth	Leverage
Mean	0.28	0.67	0.30	0.68
Small	0.55	0.67	0.34	0.65
Large	0.01	0.67	0.26	0.71
Difference	0.54	0.00	0.08	-0.06
Entrants				
Mean	1.01	0.52	1.12	0.54
Small	2.52	0.55	0.30	0.43
Large	0.20	0.50	2.34	0.72
Difference	2.32	0.05	-2.04	-0.29