A Flexible Cost Function Model with Risk

Subal C. Kumbhakar Ana Lozano-Vivas Kai Sun

Abstract

In examining bank cost efficiency in banking inclusion of risk-taking of banks is very important. In this paper we depart from the standard modeling approach and view risk intimately related to the technology. Thus, instead of controlling for risk by viewing them as covariates in the standard cost function we argue that the technology differs with risk, thereby meaning that the parameters of the parametric cost function changes with risk in a fully flexible manner. This is accomplished by viewing the parameters of the cost function as nonparametric functions of risk. We also control for country-specific effects in a fully flexible manner by using them as arguments of the nonparametric functions along with the risk variable. The resulting cost function then becomes semiparametric. The standard parametric model becomes a special case of our semiparametric model.

We use the above modeling approach for banks in the EU countries. Actually, European financial integration is seen as a stepping stone for the development of a competitive single EU market that promotes efficiency and increases consumer welfare, changing the risk profile of the European banks. Particularly, financial integration allows more risk diversification and permits banks to use more advanced risk management instruments and systems, however it has at the same time increased the probability of systematic risks. Financial integration has increased the risk of contagion and changed its nature and scope. Consequently the bank's risk seems to be an important issue to be investigated.

Keywords: Cost function; Semiparametric model; parameter heterogeneity; nonparametric model

1 Introduction

The main business a bank is to facilitate market knowledge, transaction efficiency and funding capability to customers through their services using their own balance sheet to promote those transaction and to absorb the risks associated with them. Thus, bank business is risky because in the process of providing financial services, they assume risk and then it is an essential ingredient in bank production and cost transactions. Consequently, the bank's risk is an important issue to be investigated. Since the European financial integration has lead to an important change on the risk profile of the European banks. Increased financial integration has increased the contagion of risk and changed its nature and scope. Particularly, financial integration allows more risk diversification and permits banks to use more advanced risk management instruments and systems, however it has at the same time increased the probability of systematic risks. Thus, it seems that risk is an interesting issue to be investigated in the case of European banking systems. Since risk exposure provides banks higher profit but the right management of it requires incurring in cost, it is desirable to examine bank performance and its association with risk. That is, from an economic point of view, it is desirable that not only a bank should be efficient but also that it should not, while pursuing cost minimization or profit maximization behavior, incur excessive risk-taking which may endanger the future viability of the firm. In this sense, it is necessary to measure riskiness of banks in order to be able to evaluate the bank performance correctly.

In examining bank cost efficiency in banking inclusion of risk-taking of banks is very important. In this paper we depart from the standard modeling approach and view risk intimately related to the technology. Thus, instead of controlling for risk by viewing them as covariates in the standard cost function we argue that the technology differs with risk, thereby meaning that the parameters of the parametric cost function changes with risk in a fully flexible manner. This is accomplished by viewing the parameters of the cost function as nonparametric functions of risk. We also control for country-specific effects in a fully flexible manner by using them as arguments of the nonparametric functions along with the risk variable. The resulting cost function then becomes semiparametric. The standard parametric model becomes a special case of our semiparametric model. The advantage of our semiparametric model over a standard cost function model with risk a 'shifter' is that: (1) the semiparametric model allows risk to shift the cost frontier non-neutrally, and (2) the semiparametric model estimates the impact of risk in a flexible manner, and therefore allows researcher to investigate the impact of risk for each bank during each time period.

2 Literature

The academic literature has been concentrated in analyzing the procedures of risk management and how to control it (Santomero, 1995; Mester, 1997). In these papers they introduce risk in the cost function as a covariate to control their effect on the technology (usually the cost function).

3 Econometric Model

We assume that producers use K inputs to produce Q outputs. A Cobb-Douglas cost function can be specified as:

$$\ln C = \alpha_0 + \alpha_t t + \sum_{q=1}^{Q} \alpha_q \ln Y_q + \sum_{k=1}^{K} \beta_k \ln W_k + \gamma(Z) \tag{1}$$

where C denotes total cost, t denotes a time trend, Y_q $\forall q = 1, \ldots, Q$ and W_k $\forall k = 1, \ldots, K$ represents the q-th output quantity and k-th input price, respectively, Z is the environmental factor (risk in our case). Usually $\gamma(Z)$ is specified as γZ, which means that Z neutrally shifts the cost function. The shift can be non-neutral if one specifies $\gamma(Z)$ that includes not only Z but also the interactions of Z with other variables (prices, outputs and time) in the model.

Here we consider a framework in which the Z variable plays a bigger role. That is, we allow Z to affect both the intercept and all slope coefficients in an unknown fashion, and write the cost function as:

$$\ln C = \alpha_0(Z) + \alpha_t(Z)t + \sum_{q=1}^{Q} \alpha_q(Z) \ln Y_q + \sum_{k=1}^{K} \beta_k(Z) \ln W_k, \tag{2}$$

where all the coefficients are expressed as some unknown smooth function of Z. These coefficients are therefore called smooth coefficients. The generality of the model is that Z shifts the technology in a completely flexible manner, directly through the intercept (neutral effects) and indirectly through all other variables in the model, viz., input prices, outputs and time. Imposing the restriction of homogenous of degree one in input price, viz., $\sum_k \beta_k(Z) = 1$, gives us the estimating equation:

$$\ln \widetilde{C} = \alpha_0(Z) + \alpha_t(Z)t + \sum_{q=1}^{Q} \alpha_q(Z) \ln Y_q + \sum_{k=2}^{K} \beta_k(Z) \ln \widetilde{W}_k + u, \tag{3}$$

where $\widetilde{C} = C/W_1$, $\widetilde{W}_k = W_k/W_1$ $\forall k = 1, \ldots, K$, the price of the first input is chosen as the numeraire, and u denotes the *iid* disturbance. This is the semiparametric smooth coefficient (SPSC) model first proposed by Hastie and Tibshirani (1993). Li et al. (2002) established the consistency and asymptotic normality of the SPSC estimator. The marginal effect of Z can be estimated for each bank in the data set. From this model, $\alpha_q(Z)\ \forall q = 1, \ldots, Q$, $\beta_k(Z)\ \forall k = 2, \ldots, K$, and $\alpha_t(Z)$ represents output elasticities, input price elasticities, and technical change, respectively. Returns to scale can be calculated as $1/\sum_{q=1}^{Q} \alpha_q(Z)$.

If we re-write the model (after adding the subscript i and t for observation) as:

$$\begin{aligned} \mathcal{Y}_{it} &= \alpha_0(Z_{it}) + \mathcal{X}_{it}'\Psi(Z_{it}) + u_{it} \\ &= \mathcal{W}_{it}'\Phi(Z_{it}) + u_{it} \end{aligned} \tag{4}$$

where $\mathcal{Y}_{it} = \ln \widetilde{C}_{it}$, $\mathcal{X}_{it}$ is a vector which contains all the right-hand-side regressors in (3), $\Psi(Z_{it})$ is a vector which contains all the corresponding (slope) smooth coefficients, $\mathcal{W}_{it}' = [1\ \mathcal{X}_{it}']$, and $\Phi'(Z_{it}) = [\alpha_0(Z_{it})\ \Psi'(Z_{it})]$. Following Li et al. (2002), the SPSC estimator for $\Phi(z)$ is:

$$\hat{\Phi}(z) = \left[\sum_{i=1}^{N}\sum_{t=1}^{T} \mathcal{W}_{it}\mathcal{W}_{it}' K\left(\frac{Z_{it} - z}{h}\right)\right]^{-1} \sum_{i=1}^{N}\sum_{t=1}^{T} \mathcal{W}_{it}\mathcal{Y}_{it} K\left(\frac{Z_{it} - z}{h}\right) \tag{5}$$

where N and T denotes number of countries and time periods, respectively, h is a bandwidth selected via least-squares cross-validation (Li and Racine 2010) for the Z variable and $K(\cdot)$ is the Gaussian kernel function. It differs from simple ordinary least squares (OLS) estimator only in the additional kernel function: elimination of the kernel function in (5) reduces the estimator from a smooth coefficient to its OLS counterpart. The marginal effect of z can be calculated as $\partial\mathcal{Y}_{it}/\partial z = \mathcal{W}_{it}'(\partial\hat{\Phi}(z)/\partial z)$, where

$$\frac{\partial\hat{\Phi}(z)}{\partial z} = \sum_{i=1}^{N}\sum_{t=1}^{T} \frac{\partial A_{it}(z)}{\partial z}\mathcal{Y}_{it}, \tag{6}$$

where $A_{it}(z) = A^{-1}\mathcal{W}_{it}K(\frac{Z_{it}-z}{h})$, and $A = \sum_{i=1}^{N}\sum_{t=1}^{T} \mathcal{W}_{it}\mathcal{W}_{it}'K(\frac{Z_{it}-z}{h})$.

4 Data

Empirical application Sample: Fifteen European Union countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden and the UK (**one more**). Panel Data: 7970 observations. Time period: 1993-2009.

Variables used in the cost function: banking outputs loans (Y1) and other earning assets (Y2, Total Assets-loans-physical capital). As banking inputs (prices), purchased funds and core deposits;

labor and physical capital.

For the measurement of risk, define a global measure of default risk (i.e., Z score) as

$$\text{Z score} = \frac{EA_{it} + ROA_i}{\sigma_{ROA_i}} \tag{7}$$

where EA_{it} is the average equity to total assets ratio for bank i over time, ROA_i is the average return on assets for bank i over time, and σ_{ROA_i} is the standard deviation of return on assets for each bank i over time.

In addition to the use of Z score, we also used an alternative measure of risk, defined as $P = 1/(1 + Z\ score^2)$, interpreted as a bank's probability of insolvency. An increase in P (or a decrease in Z) indicates an increase in risk. One can see that P is a normalized Z, that is, P is naturally bounded between 0 and 1, and therefore, is preferred when it comes to interpreting the results.

Table 1: **Summary Statistics of the Variables**

Symbol	Variable Name	Mean	Sd.	Min.	Max.
C	Total cost	685545.4	3123963	904	74896214
Y_1	Output 1	6296648	29734304	17.5	746347327.5
Y_2	Output 2	8015268	44451336	49.5	984785992
W_k	Capital price	9.995675	84.94299	0.005382699	4394.056537
W_l	Labor price	0.016375	0.024924	0.0000325	0.6961001
W_f	Financial funds price	0.093127	3.104442	0.00021197	275.3787879
Z	Z score	43.08872	123.3433	-69.28621858	3989.187398
P	P score	0.042855	0.126714	6.28393E-08	0.999989
t	Time trend	10.78519	4.673588	1	17

1. Total number of obs. = 7970.
2. t is calculated as $year - 1992$, where $year$ goes from 1993 to 2009.

5 Results

Table 2: Country-specific results (using P score)

Country name	Intercept	$\frac{\partial \ln \widetilde{C}}{\partial \ln \widetilde{W}_l}$	$\frac{\partial \ln \widetilde{C}}{\partial \ln \widetilde{W}_f}$	$\frac{\partial \ln \widetilde{C}}{\partial \ln Y_1}$	$\frac{\partial \ln \widetilde{C}}{\partial \ln Y_2}$	$\frac{\partial \ln \widetilde{C}}{\partial t}$	RTS	$\frac{\partial \ln \widetilde{C}}{\partial P}$
AT	3.2348	0.3746	0.4794	0.4783	0.3651	0.0119	1.1857	0.0458
	(0.1565)	(0.0189)	(0.0226)	(0.0226)	(0.0173)	(0.0005)	(0.0567)	(0.0068)
BE	1.7584	0.3059	0.5886	0.2072	0.7468	0.0075	1.0481	0.0162
	(0.0904)	(0.0154)	(0.0295)	(0.0106)	(0.0379)	(0.0004)	(0.0541)	(0.0032)
DE	3.1669	0.4649	0.3962	0.3988	0.4591	0.0202	1.1656	0.0243
	(0.0802)	(0.0115)	(0.0097)	(0.0097)	(0.0111)	(0.0005)	(0.0286)	(0.0029)
DK	1.8119	0.4896	0.4145	0.5391	0.4293	0.0007	1.0326	0.0022
	(0.0854)	(0.0226)	(0.0193)	(0.0257)	(0.0193)	(0.0000)	(0.0477)	(0.0040)
ES	3.3232	0.2647	0.6342	0.4174	0.4328	0.0199	1.1761	0.0067
	(0.1583)	(0.0126)	(0.0317)	(0.0197)	(0.0205)	(0.0010)	(0.0552)	(0.0017)
FI	3.0351	0.5396	0.3743	0.4725	0.4347	0.0065	1.1024	0.0099
	(0.3992)	(0.0705)	(0.0477)	(0.0612)	(0.0551)	(0.0008)	(0.1428)	(0.0026)
FR	2.2362	0.3119	0.5297	0.4240	0.4981	0.0052	1.0844	0.0249
	(0.0584)	(0.0080)	(0.0142)	(0.0108)	(0.0124)	(0.0002)	(0.0280)	(0.0028)
GB	3.5662	0.3228	0.5843	0.4546	0.3787	0.0093	1.2001	0.0946
	(0.1691)	(0.0152)	(0.0272)	(0.0221)	(0.0176)	(0.0004)	(0.0561)	(0.0127)
GR	2.0300	0.4909	0.4560	0.4951	0.4805	-0.0016	1.0251	0.0121
	(0.1842)	(0.0442)	(0.0417)	(0.0444)	(0.0429)	(0.0002)	(0.0925)	(0.0019)
IE	2.4641	0.5394	0.6215	0.2082	0.7608	0.1163	1.0591	-0.0608
	(0.2741)	(0.0575)	(0.0660)	(0.0227)	(0.0782)	(0.0123)	(0.1185)	(0.0236)
IT	2.7407	0.3451	0.5342	0.4333	0.4520	0.0187	1.1295	0.0112
	(0.0915)	(0.0108)	(0.0172)	(0.0146)	(0.0152)	(0.0006)	(0.0362)	(0.0024)
LU	2.1261	0.2154	0.7450	0.2160	0.7095	0.0127	1.0806	-0.0047
	(0.0703)	(0.0072)	(0.0243)	(0.0069)	(0.0237)	(0.0004)	(0.0372)	(0.0022)
NL	2.9608	0.1682	0.6481	0.1238	0.6577	0.0553	1.2797	0.0648
	(0.2754)	(0.0154)	(0.0590)	(0.0116)	(0.0604)	(0.0051)	(0.1190)	(0.0154)
PT	1.7807	0.3299	0.4700	0.5992	0.3650	-0.0239	1.0371	0.0114
	(0.1362)	(0.0269)	(0.0354)	(0.0466)	(0.0290)	(0.0019)	(0.0849)	(0.0016)
SE	3.1081	0.4778	0.4161	0.4288	0.4621	0.0020	1.1224	-0.0142
	(0.2298)	(0.0380)	(0.0318)	(0.0332)	(0.0343)	(0.0002)	(0.0849)	(0.0021)

1. Mean estimates are reported.
2. The numbers in the parentheses are standard errors.
3. P score is used as measure of risk.

Table 3: Size-specific results (using P score)

Size	Intercept	$\frac{\partial \ln \widetilde{C}}{\partial \ln \widetilde{W}_l}$	$\frac{\partial \ln \widetilde{C}}{\partial \ln \widetilde{W}_f}$	$\frac{\partial \ln \widetilde{C}}{\partial \ln Y_1}$	$\frac{\partial \ln \widetilde{C}}{\partial \ln Y_2}$	$\frac{\partial \ln \widetilde{C}}{\partial t}$	RTS	$\frac{\partial \ln \widetilde{C}}{\partial P}$
$\ln(asset) \leq Q_1$	2.7147	0.3719	0.5040	0.4086	0.4808	0.0126	1.1272	-0.0001
	(0.0605)	(0.0085)	(0.0124)	(0.0097)	(0.0111)	(0.0004)	(0.0253)	(0.0022)
$Q_1 < \ln(asset) \leq Q_2$	2.6529	0.3571	0.5264	0.3913	0.5027	0.0127	1.1215	0.0166
	(0.0608)	(0.0084)	(0.0125)	(0.0089)	(0.0117)	(0.0004)	(0.0249)	(0.0019)
$Q2 < \ln(asset) \leq Q3$	2.6053	0.3482	0.5363	0.3870	0.5103	0.0136	1.1173	0.0209
	(0.0613)	(0.0079)	(0.0123)	(0.0090)	(0.0111)	(0.0005)	(0.0249)	(0.0023)
$\ln(asset) > Q_3$	2.6117	0.3520	0.5375	0.3926	0.5067	0.0144	1.1156	0.0444
	(0.0611)	(0.0084)	(0.0122)	(0.0090)	(0.0112)	(0.0005)	(0.0257)	(0.0037)

1. Mean estimates are reported.
2. The numbers in the parentheses are standard errors.
3. P score is used as measure of risk.

Table 4: Country-specific results (using Z score)

Country name	Intercept	$\frac{\partial \ln \widetilde{C}}{\partial \ln \widetilde{W}_l}$	$\frac{\partial \ln \widetilde{C}}{\partial \ln \widetilde{W}_f}$	$\frac{\partial \ln \widetilde{C}}{\partial \ln Y_1}$	$\frac{\partial \ln \widetilde{C}}{\partial \ln Y_2}$	$\frac{\partial \ln \widetilde{C}}{\partial t}$	RTS	$\frac{\partial \ln \widetilde{C}}{\partial Z}$
AT	3.0121	0.2962	0.5529	0.4902	0.3671	0.0105	1.1690	-0.0009
	(0.01064)	(0.00136)	(0.00773)	(0.00190)	(0.00054)	(0.00003)	(0.00120)	(0.00017)
BE	2.1038	0.2847	0.5998	0.2502	0.6752	0.0139	1.0806	-0.0004
	(0.00054)	(0.00004)	(0.00056)	(0.00032)	(0.00121)	(0.00002)	(0.00007)	(0.00008)
DE	3.0429	0.3928	0.4545	0.4075	0.4496	0.0126	1.1587	-0.0006
	(0.00165)	(0.00028)	(0.00103)	(0.00023)	(0.00023)	(0.00001)	(0.00014)	(0.00008)
DK	2.3755	0.3979	0.5019	0.4188	0.5077	0.0044	1.0807	-0.0002
	(0.00087)	(0.00006)	(0.00047)	(0.00097)	(0.00196)	(0.00009)	(0.00012)	(0.00009)
ES	3.3200	0.2344	0.6843	0.4454	0.4068	0.0163	1.1704	-0.0007
	(0.02346)	(0.00015)	(0.00370)	(0.00103)	(0.00037)	(0.00002)	(0.00119)	(0.00012)
FI	2.8685	0.4165	0.4577	0.3993	0.4909	0.0090	1.1210	-0.0021
	(0.00502)	(0.00164)	(0.00479)	(0.02774)	(0.00219)	(0.00008)	(0.00021)	(0.00029)
FR	2.2896	0.2677	0.5745	0.4278	0.4833	0.0106	1.0969	-0.0008
	(0.00371)	(0.00029)	(0.00188)	(0.00019)	(0.00018)	(0.00004)	(0.00027)	(0.00007)
GB	3.5773	0.2827	0.6400	0.4234	0.4148	0.0068	1.2045	-0.0030
	(0.00358)	(0.00248)	(0.00538)	(0.00082)	(0.00432)	(0.00018)	(0.00020)	(0.00023)
GR	2.5689	0.3381	0.5484	0.4102	0.4900	0.0188	1.1074	-0.0014
	(0.00223)	(0.00242)	(0.00471)	(0.00073)	(0.00040)	(0.00016)	(0.00084)	(0.00022)
IE	3.3874	0.2643	0.6603	0.3325	0.5112	0.0248	1.1877	0.0003
	(0.13982)	(0.00019)	(0.00929)	(0.00869)	(0.00016)	(0.00023)	(0.00921)	(0.00029)
IT	2.8722	0.3078	0.5738	0.4477	0.4273	0.0182	1.1389	-0.0007
	(0.00040)	(0.00049)	(0.00219)	(0.00216)	(0.00025)	(0.00001)	(0.00005)	(0.00008)
LU	2.4384	0.2248	0.7393	0.2529	0.6579	0.0146	1.0975	-0.0002
	(0.00041)	(0.00017)	(0.00014)	(0.00021)	(0.00082)	(0.00004)	(0.00002)	(0.00006)
NL	3.2632	0.2165	0.6210	0.2089	0.5897	0.0375	1.2511	-0.0041
	(0.00661)	(0.00198)	(0.01543)	(0.00015)	(0.00141)	(0.00058)	(0.00349)	(0.00044)
PT	2.6604	0.2992	0.5872	0.4629	0.4429	0.0040	1.1063	-0.0010
	(0.03420)	(0.00055)	(0.01609)	(0.00105)	(0.00289)	(0.00032)	(0.00267)	(0.00016)
SE	3.1510	0.3839	0.5296	0.3964	0.4838	0.0064	1.1349	0.0012
	(0.01619)	(0.00127)	(0.00280)	(0.00145)	(0.00053)	(0.00009)	(0.00067)	(0.00020)

1. Median estimates are reported, because outlier estimates distort the mean.
2. The numbers in the parentheses are standard errors.
3. Z score is used as measure of risk.

Table 5: Size-specific results (using Z score)

Size	Intercept	$\frac{\partial \ln \widetilde{C}}{\partial \ln \widetilde{W}_l}$	$\frac{\partial \ln \widetilde{C}}{\partial \ln \widetilde{W}_f}$	$\frac{\partial \ln \widetilde{C}}{\partial \ln Y_1}$	$\frac{\partial \ln \widetilde{C}}{\partial \ln Y_2}$	$\frac{\partial \ln \widetilde{C}}{\partial t}$	RTS	$\frac{\partial \ln \widetilde{C}}{\partial Z}$
$\ln(asset) \leq Q_1$	2.8597	0.3764	0.5036	0.4115	0.4699	0.0129	1.1345	0.0001
	(0.00858)	(0.00116)	(0.00173)	(0.00086)	(0.00039)	(0.00001)	(0.00035)	(0.00004)
$Q_1 < \ln(asset) \leq Q_2$	2.8414	0.3118	0.5495	0.4117	0.4742	0.0125	1.1307	-0.0007
	(0.00270)	(0.00203)	(0.00234)	(0.00108)	(0.00092)	(0.00001)	(0.00016)	(0.00005)
$Q2 < \ln(asset) \leq Q3$	2.6915	0.2887	0.5773	0.4127	0.4770	0.0131	1.1185	-0.0006
	(0.00986)	(0.00087)	(0.00267)	(0.00101)	(0.00109)	(0.00001)	(0.00011)	(0.00005)
$\ln(asset) > Q_3$	2.8569	0.2892	0.5854	0.4139	0.4762	0.0136	1.1301	-0.0015
	(0.00216)	(0.00106)	(0.00341)	(0.00104)	(0.00098)	(0.00007)	(0.00016)	(0.00006)

1. Median estimates are reported, because outlier estimates distort the mean.
2. The numbers in the parentheses are standard errors.
3. Z score is used as measure of risk.