Innovation and Productivity in Firm Level: an Empirical Study of China Textile Industry

Feng et al.

Abstract: …

JEL Classification: D21 L67 O31

Key Words: CDM Model, R&D, New Product, Productivity

1. Introduction

Innovation is a key word in moving into the knowledge-based economy not only in developed US and EU countries' development, but also in China's reform and development. After 30 years reform, China gains a long term of fast development and becomes a "world factory". In the new century, China is seeking a new developing approach to improve productivity, save energy and resources, but keeps the fast developing speed and aims to be the world manufacturing center at the same time. The practicable way is to establish a knowledge-based economy and make innovation a main factor in research and development, production and management, especially in manufacturing sectors in firm level. That is the Sustainable Development Strategy and Innovation-oriented Country Strategy came into being in the national level.

Textile industry is the largest manufacturing sector in China in number of firms and labour force. It is also a competitive sector of China in the world market and a sector easy to raise trade conflicts. It has sufficient power to influence employment market in China and textile product market in the world. By using the firm level census data, the sample size is complicated enough for an empirical study concentrate on a special sector. We will use CDM model to understand the relations between innovation input, innovation output and productivity, by using textile sector as an example.

The paper organized as follows. Section 2 introduces CDM model and its estimation method in this paper. Section 3 describes data and variables selection. The results are presented in section 4 and conclusions are drawn in section 5.

2. Models and Estimation

CDM model gives us a systematic understanding of innovation path in production. It has 3 steps and 4 equations written as follows, with i=1,⋯,n index firms and t index year. Vector x series are explanatory variables, vector b series are parameters and vector u series are error terms.

Innovation input:

$$brd_{(t-1)i} = X_{0(t-1)i}b_0 + u_{0i} \quad (1a)$$

$$lrdps_{(t-1)i} = X_{1(t-1)i}b_1 + u_{1i} \quad (1b)$$

Innovation output:

$$bnp_{ti} = \alpha * \widehat{lrdps}_{(t-1)i} + X_{2ti}b_2 + u_{2i} \quad (2)$$

Innovation performance:

$$lp_{ti} = \gamma * \widehat{bnp}_{ti} + X_{3ti}b_3 + u_{3i} \quad (3)$$

Step 1 explains innovation input with 2 equations shown as (1a) and (1b). The first equation (1a) is a probit model as a selection equation to understand firms' decision of input or not in innovation. The second equation (1b) is a Tobit model to explain how much they would like to expend in innovation. Step 2 with equation (2) is a knowledge production function, explains innovation input and its influence on innovation output. Step 3 with equation (3) is an extended Cobb-Douglas Production Function to explain innovation output and its influence on productivity.

We use Heckman procedure in STATA to estimate the first 2 equations with 1 year earlier than the following 2 steps. Innovation input is measured by binary variable brd to explain whether firms' are input or not in probit model, and by R&D expenses per sales after logarithm ($lrdps$) in Tobit model to explain how much they would like

to innovate. Innovation output is measured by a binary variable bnp to explain whether firms have new product output or not. Labour Productivity (lp) is an explained variable in the last step. In order to involve all the firms in the estimation, under the hypothesis of all firms prefer some innovative effort, we use predicted value of innovation input ($\widehat{lrdps}$) in equation 2 and predicted value of innovation output ($\widehat{bnp}$) in equation 3. The last 2 steps estimated separately.

3. Data and Variable selection

The data comes from the yearly industrial survey organized by China National Bureau of Statistics. Infect, it is a yearly census of all state-owned firms, and non-state-owned firms above designated size[1]. The criteria are all hard data and most of them from the accounting form of the enterprise.

For the original data, we delete those firms whose employees are less than 10, sales of products are less than RMB 500 million, or value added is less then 0. Then, the growth rate of sales, labour and capital of each firm are calculated. Firms with all the 3 growth rates are between each 2.5 and 97.5 percentile are kept in the modeling. At last, we get 13366 balanced firms data in each of the 2 years, 2005 and 2006.

Three groups of variables are selected to establish CDM model. The first group is innovation variables, with R&D expenses per sales (in log) and a group of binaries to measure R&D expenses, new product, export, advertisement, labour training expenses and subsidy income. The second group is basic variables in production function, include productivity, labour and ratio of capital per employee (all in logs). The third group is extended variables to measure market size by sales ratio in 4-digit sub-sectors[2], firm size by 4 dummies, categories of capital hold by 5 capital ratios in

[1] The designated size means Sales of Products is higher than RMB 5 million (EUR 500 thousand)
[2] It is the smallest category in China Industry Standard.

total capital hold. All variables are explained in appendix table 1. Influence of price change in equation (2) has been eliminated when calculate the R&D input per sales.

Table 1 gives the basic description of variables in each corresponding year.

[Table 1-1 & 1-2 here]

4. Result

Innovation input model in table 2 shows firms market size is only significant in selection model with a small coefficient. Large firms with employees higher than 250 has positive trend in innovation input. Advertisement and subsidy are positive in both equations, export and labour training are only positive in selection equation. Foreign capital hold has a significant negative influence in innovation input.

[Table 2 here]

Knowledge production function in table 3 shows that R&D expenses and continual innovation input are significant positive in improving innovation output. The top 2 kinds of large firms still positive in innovation output. Other innovation behaviors like advertisement, labour training are positive in improve new product. All kinds of capital hold except state capital are negative in knowledge production.

[Table 3 here]

Innovation output improves productivity and the influence become much greater when more explanatory variables are added (shown in table 4). Foreign firms have high level of productivity though they do not favour in innovation input. (There are some problems in this equation makes the coefficient unstable and some parameter difficult to explain.)

[Table 4 here]

5. Conclusion (left to be added)

Appendix Table 1: Variable Definitions

Variable Name	Variable Code	Explanation
Basic (dm gives firm id, and t is time)		
Productivity	**lp**	Value Added per employee (in log.)
Physical Capital	**lcdl**	Total assets per employee (in log.)
Labour	**lca**	Number of employees (in log.)
Innovation		
R&D Expenses (RDE)	**brd**	brd=1 if R&D expenses >0 (Binary)
	crd	crd(t)=1 if brd(t)=brd(t-1)=1
	lrdps	R&D expenses per sale (in log.)
Value of New Product (NPV)	**bnp**	bnp=1 if Values of new products >0 (Binary)
	cnp	cnp(t)=1 if bnp(t)=bnp(t-1)=1
Value of Export	**bex**	bnp=1 if Value of export >0 (Binary)
Labour Training	**blt**	bnp=1 if Labour training expenses >0 (Binary)
Advertisement	**bad**	bad=1 if advertisement expenses >0 (Binary)
Subsidy Income	**bsu**	bsu=1 if subsidy income >0 (Binary)
Others		
Market Size	**lsts**	sales divided by total sales in 4-digit sub-sector (in log.)
Firm Size	**ss**	small firms, L>=50 & L <100 (Dummy)
	sm	middle firms, L>=100 & L <250 (Dummy)
	sl	large firms, L>=250 & L <1000 (Dummy)
	sh	huge firms, L>=1000 (Dummy)
Capital Hold	**rsc**	ratio of State Capital in total capital hold (It is not included in model to avoid collinearity)
	rcl	ratio of Collective Capital in total capital hold
	rcp	ratio of Corporate Capital in total capital hold
	ric	ratio of Individual Capital in total capital hold
	rhc	ratio of Hong Kong, Macau and Taiwan Capital in total capital hold
	rfc	ratio of Foreign Capital in total capital hold

Table 1-1: Description of variables (2005)

Variable	Obs	Mean	Std. Dev.	Min	Max
lrdps	753	-6.192491	1.70121	-12.583760	-1.516313
brd	13366	.0563370	.2305799	0	1
bnp	13366	.0711507	.2570860	0	1
lla	13366	4.999853	1.059012	2.302585	11.45226
lsts	13366	-8.350879	1.56368	-11.543650	-1.198565
bad	13366	.1182852	.3229576	0	1
bex	13366	.4092473	.4917134	0	1
blt	13366	.3989226	.4896951	0	1
bsu	13366	.1557684	.3626492	0	1
rsc	13346	.0161535	.1149224	0	1
rcl	13346	.0393750	.1797762	0	1
rcp	13346	.1933488	.3650897	0	1
ric	13346	.5800121	.4715853	0	1
rhc	13346	.0995730	.2792388	0	1
rfc	13346	.0715376	.2368037	0	1
ss	13366	.2400120	.4271064	0	1
sm	13366	.3483466	.4764643	0	1
sl	13366	.2290887	.4202622	0	1
sh	13366	.0522969	.2226334	0	1

Table 1-2: Description of variables (2006)

Variable	Obs	Mean	Std. Dev.	Min	Max
lp	13366	3.993733	.8482531	3.355735	7.794494
lcdl	13366	4.888228	.9085365	1.003318	9.036594
lla	13366	5.010757	1.055825	2.302585	11.81627
lsts	13366	8.414796	1.57503	11.75915	1.439342
bnp	13366	.0834954	.27664	0	1
brd	13366	.0611252	.2395689	0	1
bad	13365	.1122334	.3156653	0	1
bex	13366	.4026635	.4904524	0	1
blt	13366	.4065539	.4912086	0	1
bsu	13366	.1417776	.348835	0	1
rsc	13347	.0141208	.1065838	0	1
rcl	13347	.0356047	.1719057	0	1
rcp	13347	.1964038	.367073	0	1
ric	13347	.5835795	.4714819	0	1
rhc	13347	.1012082	.2813447	0	1
rfc	13347	.069083	.2322368	0	1
ss	13366	.2366452	.4250386	0	1
sm	13366	.35104	.4773133	0	1
sl	13366	.2324555	.4224136	0	1
sh	13366	.0524465	.222934	0	1
crd	13366	.0350142	.1838226	0	1
cnp	13366	.0489301	.21573	0	1

Table 2: Innovation Input Equations

Heckman selection model — Number of obs = 13346
(regression model with sample selection) — Censored obs = 12593
Uncensored obs = 753
Wald chi2(10) = 36.27
Log likelihood = 3825.57 — Prob > chi2 = 0.0001

		Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
lrdps	lsts	.0647845	.0457859	1.41	0.157	.0249543	.1545233
	bad	.6871064	.2481048	2.77	0.006	.20083	1.173383
	bex	.1506366	.1450999	1.04	0.299	.1337539	.4350271
	blt	.3245575	.1703547	1.91	0.057	.0093315	.6584464
	bsu	.4692914	.1510582	3.11	0.002	.1732228	.76536
	rcl	1.1279150	.5517674	2.04	0.041	2.20936	.046471
	rcp	.1855512	.4035291	0.46	0.646	.9764537	.6053512
	ric	.0367955	.3909978	0.09	0.925	.8031371	.729546
	rhc	1.3716590	.4934324	2.78	0.005	2.338768	.404549
	rfc	.5264627	.4818374	1.09	0.275	1.470847	.4179213
	_cons	8.402713	.6574594	12.78	0.000	9.69131	7.114116
brd	lsts	.0517031	.0138512	3.73	0.000	.0245552	.078851
	bad	.9144718	.0447677	20.43	0.000	.8267287	1.002215
	bex	.0879561	.0426977	2.06	0.039	.0042701	.1716421
	blt	.3790725	.0411258	9.22	0.000	.2984674	.4596776
	bsu	.1279045	.0485428	2.63	0.008	.0327623	.2230467
	rcl	.4187141	.1744875	2.40	0.016	.7607034	.0767248
	rcp	.2201179	.1393378	1.58	0.114	.493215	.0529791
	ric	.2488733	.1348894	1.85	0.065	.5132516	.0155049
	rhc	.6486718	.1540348	4.21	0.000	.9505744	.3467692
	rfc	.4543042	.1559193	2.91	0.004	.7599005	.1487079
	ss	.0975266	.0763295	1.28	0.201	.0520765	.2471296
	sm	.0627930	.0731753	0.86	0.391	.080628	.2062141
	sl	.2952740	.0753592	3.92	0.000	.1475726	.4429754
	sh	.6203490	.0928410	6.68	0.000	.438384	.802314
	_cons	1.5320530	.1967128	7.79	0.000	1.917604	1.146503
/athrho		.7506483	.1484327	5.06	0.000	.4597256	1.041571
/lnsigma		.7006939	.0729117	9.61	0.000	.5577895	.8435983
rho		.6355356	.0884799			.4298606	.7785077
sigma		2.015151	.1469281			1.746807	2.324717
lambda		1.2807	.2676991			.7560194	1.80538

LR test of indep. eqns. (rho = 0): chi2(1) = 15.92 Prob > chi2 = 0.0001

Table 3: Innovation Onput Equations

Probit regression — Number of obs = 13329; Wald chi2(16) = 903.66; Prob > chi2 = 0.0000; Log pseudolikelihood = 3321.5619; Pseudo R2 = 0.1303

bnp	Coef.	Robust Std. Err.	z	P>z	[95% Conf.	Interval]
lrdps_ha	.1159349	.0470913	2.46	0.014	.0236377	.208232
t	.6822403	.0692105	9.86	0.000	.5465901	.8178904
crd	.02087	.0131246	1.59	0.112	.0048539	.0465938
lsts	.161016	.0523489	3.08	0.002	.058414	.2636179
bad	.6495414	.0388261	16.73	0.000	.5734437	.7256392
bex	.0779173	.0364834	2.14	0.033	.006411	.1494235
blt	.1530781	.0455843	3.36	0.001	.0637344	.2424217
bsu	.5464239	.1574037	3.47	0.001	.8549295	.2379184
rcl	.5454333	.121466	4.49	0.000	.7835024	.3073642
rcp	.4718479	.1168353	4.04	0.000	.700841	.2428549
ric	.923745	.1463907	6.31	0.000	1.210666	.6368244
rhc	.925036	.1396622	6.62	0.000	1.198769	.6513031
rfc	.0294276	.0649637	0.45	0.651	.0978989	.156754
ss	.0232243	.0620966	0.37	0.708	.1449313	.0984827
sm	.1397726	.0652762	2.14	0.032	.0118337	.2677116
sl	.3731534	.0844403	4.42	0.000	.2076534	.5386533
sh	.1789908	.4212262	0.42	0.671	1.004579	.6465973
_cons						

Table 4-1: Production Function Equations

Linear regression — Number of obs = 13329; F(3, 13325) = 1382.83; Prob > F = 0.0000; Rsquared = 0.2457; Root MSE = .73623

lp	Coef.	Robust Std. Err.	t	P>t	[95% Conf.	Interval]
lcdl	.3959411	.0073809	53.64	0.000	.3814735	.4104088
lla	.1568217	.0069088	22.70	0.000	.1703639	.1432795
bnp_hat	.0273506	.0148617	1.84	0.066	.0017805	.0564817
_cons	2.886246	.0673009	42.89	0.000	2.754327	3.018165

Table 4-2: Extended Production Function Equations

Linear regression

Number of obs = 13329
F(17, 13311) = 272.22
Prob > F = 0.0000
Rsquared = 0.2674
Root MSE = .72594

lp	Coef.	Robust Std. Err.	t	P>t	[95% Conf.	Interval]
lcdl	.3866049	.0077454	49.91	0.000	.3714228	.4017869
lla	.2052794	.0213851	9.60	0.000	.2471971	.1633616
bnp_hat	.5782968	.04487	12.89	0.000	.4903452	.6662483
bex	.4315615	.033697	12.81	0.000	.4976124	.3655106
blt	.1287495	.0143668	8.96	0.000	.1569105	.1005885
bad	.1132972	.0230832	4.91	0.000	.1585435	.0680508
bsu	.1976994	.0191035	10.35	0.000	.2351449	.1602539
rcl	.8693283	.088791	9.79	0.000	.6952853	1.043371
rcp	.9130156	.0752411	12.13	0.000	.7655323	1.060499
ric	.7692075	.0731188	10.52	0.000	.6258841	.9125308
rhc	1.031159	.0862303	11.96	0.000	.8621351	1.200182
rfc	1.109535	.085677	12.95	0.000	.9415954	1.277474
cnp	.1288892	.0348101	3.70	0.000	.0606564	.197122
ss	.0527295	.0285252	1.85	0.065	.1086429	.0031839
sm	.0198887	.0395798	0.50	0.615	.0576933	.0974707
sl	.0180281	.0583189	0.31	0.757	.1323415	.0962852
sh	.0097383	.0893433	0.11	0.913	.1653873	.1848639
_cons	3.449916	.1445934	23.86	0.000	3.166493	3.73334

References (left to be added)