

ONLINE APPENDIX

Appendix A: Database Construction Procedures, Summary Statistics, Correlation Matrix, PSM Tests, and Control Variables

Database Construction Procedures

To obtain unique identifiers for entrepreneurs/firms in the Chinese Private Entrepreneurs Survey (CPES) Data from the Chinese Academy of Social Science (CASS), one approach is to use initial conditions and other fixed characteristics. Same firms have identical values for these variables, and this facilitates identifying the common firms over time. These identifying variables include founding conditions such as initial sources of funding, initial financial indicators (e.g., total assets, sales, registered capital, taxes, and fees), and initial employee makeup (e.g., number of technicians, managers, and workers), all of which are reported in different survey waves. This approach has been confirmed by data providers as an accurate way to identify firms and has been used by other researchers. These are the basic steps of this procedure:

1. Assuming a researcher is working from the individual raw data files for each year, s/he would first need to clean these files, particularly making variable names consistent. Generating a codebook to indicate the common names and available years of each variable is a helpful first step. The researcher needs to attend to this closely as order of questions, naming conventions, and other details in the survey questionnaires were not consistent over time. Generate the line number of each firm in the original dataset, and this will be the year-ID of that firm in that year. After this, the researcher can pool the original datasets together to get a full sample encompassing all the available years.
2. To get the dictionary linking year, year-IDs and the unique-IDs, merge pairs of datasets from survey waves by relying on the codebook generated from the step above, which provides availability of these identifying variables. For instance, the researcher can start by matching 1993 and 1995 data with the available identifying variables, e.g., initial investments variables. Then 1993 with 1997, 2000, 2002, and so on; 1995 with 1997, 2000, 2002, and so on; 1997 with 2000, 2002, and so on; and so on. For each pair of two waves, select the generic identifying variables that are available—such as initial sources of funding—first, and if there are too many missing values and/or inconsistencies (e.g., in terms of rounding and other reporting errors), then use other identifying variables to match the two waves. Due to different sources of inconsistency such as input or reporting errors, frequently manual inspection and human judgment is necessary to determine whether the two observations indeed match. This step results in pairs of matches of year-IDs, e.g., IDs in 1993 to be linked to IDs in 1995.

3. Regarding the unique-IDs, the researcher can set 1993 as the benchmark and begin with that year, then unique-IDs would be equal to the year-IDs or the line number in that year—1993 in this example. Find out which lines can be matched with 1993's lines in the 1995 dataset and add that corresponding unique-ID to the year-ID in 1995. Continue in this fashion, add unique-IDs for the 1997 to 2012 data. Then continue year by year and assign unique-IDs to the unmatched data. For instance, if the first unmatched line in 1995 is line (year-ID) 4, continue numbering that line based on the next number in the unique-ID sequence. For example, if the largest unique-ID number for 1993 is 1000, then assign 1001 to the line 4 of 1995 and continue until every unmatched line has a unique-ID. Then use the unmatched 1995 lines with the newly assigned unique-IDs to match 1997, 2000, and so on, continuing in this fashion to get a year-ID-line dictionary. Note that the researcher would need to cross-check the match over time after the series of dyadic matches, i.e., if line 5 in 1995 and line 9 in 1997 are both matched to line 3 in 1993, then the first two should be identical in other identifying variables—if available. Again, this process involves manual inspection and human judgment in case of different sources of inconsistency such as rounding issues of decimal places (sometimes different waves report different decimal levels), missing values (and thus need other identifying variables), and potential input or reporting errors.
4. Finally, generate a dataset containing year-ID-line for all years by pooling them together. Then merge the pooled year-ID-line with the entire sample (from step 1) to create an overall database that includes unique identifiers.

Table A1. Summary Statistics and Correlation Matrix (N = 20,564)*

Variable	Mean	S.D.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Inward internationalization (0/1)	.15	.36															
2. Inward internationalization (foreign investment/total)	.11	.12	.63														
3. Outward internationalization (0/1)	.18	.38	.14	.14													
4. Outward internationalization (overseas assets/total)	.12	.48	.12	.11	.43												
5. Ideological imprint	.21	.41	-.18	-.15	-.16	-.19											
6. Political involvement	.30	.46	.04	.07	.01	.05	-.06										
7. Government appropriation	.13	.10	.05	.15	-.07	-.02	.03	.00									
8. Social network	.69	.46	.04	.07	.05	-.01	-.02	.10	.02								
9. Post 2001	.66	.47	.07	.09	.14	.02	.09	-.07	-.23	-.04							
10. Regional FDI intensity	.25	.28	.10	.20	.08	.00	.08	-.06	-.05	-.05	.49						
11. Age	45.29	8.71	.01	.01	.02	.01	.02	.04	-.08	-.01	.14	.08					
12. Educational attainment	12.89	1.62	-.03	-.02	-.02	.00	-.02	.06	.05	-.01	-.32	-.17	-.16				
13. Current communist ideology	5.53	2.52	-.01	-.01	.01	-.01	.34	-.24	.01	-.02	.06	.00	-.13	-.02			
14. Foreign experience	.21	.41	.03	.04	.05	.01	.05	.29	-.09	.09	.33	-.07	.00	-.14	-.01		
15. Government work experience	.09	.28	.02	.02	.04	.00	.05	-.08	-.05	-.04	.18	.11	.05	-.04	.02	.00	
16. Firm size	3.61	1.69	.02	.02	.02	.01	.02	.20	-.06	.03	.03	.04	.11	.07	-.24	.04	.00
17. Financial leverage	.13	.22	.01	.02	.00	-.01	.01	.05	.06	.02	.03	-.04	.04	-.05	-.05	.12	-.01
18. Firm age	7.89	5.59	.05	.05	.11	.01	.10	-.21	-.22	-.19	.49	.45	.15	-.18	.06	-.05	.22
19. Industry average (inward internationalization 0/1)	.15	.04	.06	.05	.11	.02	.00	.07	-.20	.03	.41	.35	.11	-.23	.01	.28	.09
20. Industry average (inward internationalization, continuous)	.11	.04	.95	.62	.04	.01	-.06	.03	.05	.03	.10	.12	.01	-.04	.00	.06	.03
21. Industry average (outward internationalization 0/1)	.18	.09	.05	.04	.24	.05	.00	.00	-.27	.06	.49	.25	.09	-.09	.04	.18	.10
22. Industry average (outward internationalization, continuous)	.12	.06	.02	.02	.69	.22	-.07	.07	-.05	.01	.05	.05	.00	.01	-.01	.03	.00
23. Manufacturing firm	.22	.23	.00	-.01	.01	.01	-.01	.00	-.01	-.05	.02	.04	.03	-.03	-.04	-.08	.01
24. Poor Internet coverage	.11	.14	.00	.00	-.02	.01	.00	.00	.00	-.01	.00	.00	.00	.00	.00	.01	-.01
25. Density of CPC membership	.06	.06	.01	.00	-.01	.00	.00	.01	-.01	.01	.00	.01	.00	.00	.00	.00	.00
26. GDP per capita (logged)	7.25	4.21	.03	.02	.09	.04	.01	-.10	-.21	-.11	.29	.46	.08	-.03	.01	-.42	.09
27. Population growth	.04	.04	-.01	-.01	.02	.03	-.01	.05	-.08	.15	-.15	-.04	-.04	.07	.02	-.18	-.08
28. Institutional development	6.76	3.03	.03	.03	.10	.02	.07	-.08	-.19	.07	.42	.33	.06	-.12	.06	.05	.09
29. R&D investment (0/1)	.37	.48	.02	.02	.05	.01	.06	.11	-.12	.06	.08	.04	.08	.08	-.09	.07	.00
30. Firm performance	.13	.28	-.03	-.02	-.07	-.02	-.09	.00	.35	-.06	-.25	-.13	-.04	.04	-.01	-.10	-.04
31. Industry competitiveness	.93	.12	-.01	.00	-.03	-.03	.01	.03	.04	.05	-.09	-.07	-.01	-.06	-.01	.08	-.04
32. Industry growth	.00	.64	.00	.00	.00	.00	.00	.00	.02	-.01	.00	.00	.01	.01	.01	.00	.00
Variable	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
17. Financial leverage	.07																
18. Firm age	.02	-.02															
19. Industry average (inward internationalization 0/1)	.08	.01	.42														
20. Industry average (inward internationalization, continuous)	.02	.02	.37	.62													
21. Industry average (outward internationalization 0/1)	.05	.02	.32	.20	.16												
22. Industry average (outward internationalization, continuous)	.02	.00	.04	.18	.19	.37											
23. Manufacturing firm	.05	-.01	.07	.02	-.01	.00	.03										
24. Poor Internet coverage	.01	.00	.00	.00	-.01	.00	-.01	.01									
25. Density of CPC membership	.01	.01	.00	.01	.00	.00	.00	.00	.00								
26. GDP per capita (logged)	.03	-.16	.48	.27	.02	.32	.11	.09	.00	.01							
27. Population growth	-.02	-.10	-.17	-.02	-.02	.09	.06	-.03	.00	-.01	.47						
28. Institutional development	-.01	-.03	.37	.29	.05	.37	.04	-.01	.01	.41	.11	.11					
29. R&D investment (0/1)	.26	.07	.07	.09	.02	.15	.03	.01	.00	.00	.15	.14	.15				
30. Firm performance	-.06	.15	-.23	-.19	-.04	-.26	-.04	.01	.00	-.01	-.35	-.25	-.30	-.18			
31. Industry competitiveness	.02	.04	-.19	-.05	.00	-.10	-.06	-.09	-.01	-.01	-.21	-.01	-.06	.00	.06		
32. Industry growth	.01	.01	-.01	.01	.00	.00	.00	.01	-.01	.00	-.04	-.03	-.03	-.01	.06	.01	

* Pearson correlation tests are used if both variables are continuous, and Spearman rank tests are employed otherwise. Coefficients of correlations over .01 are significant at the 1% level.

Table A2. Logit Regression Results, Pre- and Post-match Differences, and Percentage Bias Reduction*

Variable	Pre-match	Post-match		Bias (%)	
	1 Results	2 <i>p</i> -value of differences	3 Reduction	4 Post-match	
Age	.015** (.001)	.324	388.6	–.511	
Educational attainment	.067** (.007)	.198	239.9	1.19	
Current communist ideology	.024** (.004)	.277	421.9	1.24	
Government work experience	.197** (.006)	.347	226.1	.91	
Foreign experience	–.870** (.038)	.28	492.1	.8	
Firm size	.054** (.007)	.152	354.7	1.5	
Financial leverage	.105* (.051)	.202	233.7	–.66	
Firm age	–.019** (.000)	.394	297.4	1.6	
Manufacturing firm	–.261** (.036)	.491	483.6	.4	
Industry average of inward Internationalization (0/1)	–.314** (.024)	.458	375.4	–.83	
Industry average of outward internationalization (0/1)	–.578** (.044)	.42	329.4	–.44	
GDP per capita (logged)	.177** (.006)	.364	188.0	–1.4	
Population growth	–.031** (.004)	.283	274.6	–.46	
Institutional development	–.017** (.004)	.376	135.7	–1.7	
Poor Internet coverage	–.033** (.010)	.321	138.9	.2	
Density of CPC membership	–.077* (.039)	.315	149.6	.14	
Political involvement	.168** (.029)	.219	133.8	1.36	
Government appropriation	–.576** (.132)	.285	178.8	1.33	
Social network	.508** (.028)	.154	237.1	–1.35	
Post 1999	.247** (.047)	.248	364.7	–.94	
Regional FDI intensity	–.004** (.001)	.232	269.5	–1	
Number of observations	20,564				
Goodness of fit (pseudo R ²)	.270				

* $p < .10$; * $p < .05$; ** $p < .01$; intercept not reported to save space.

* The pseudo R² of the logit model diminished to less than .001 for the matched sample, and all variables are insignificant at the 10% level.

Figure A1. Graphical illustration of PSM matching quality.

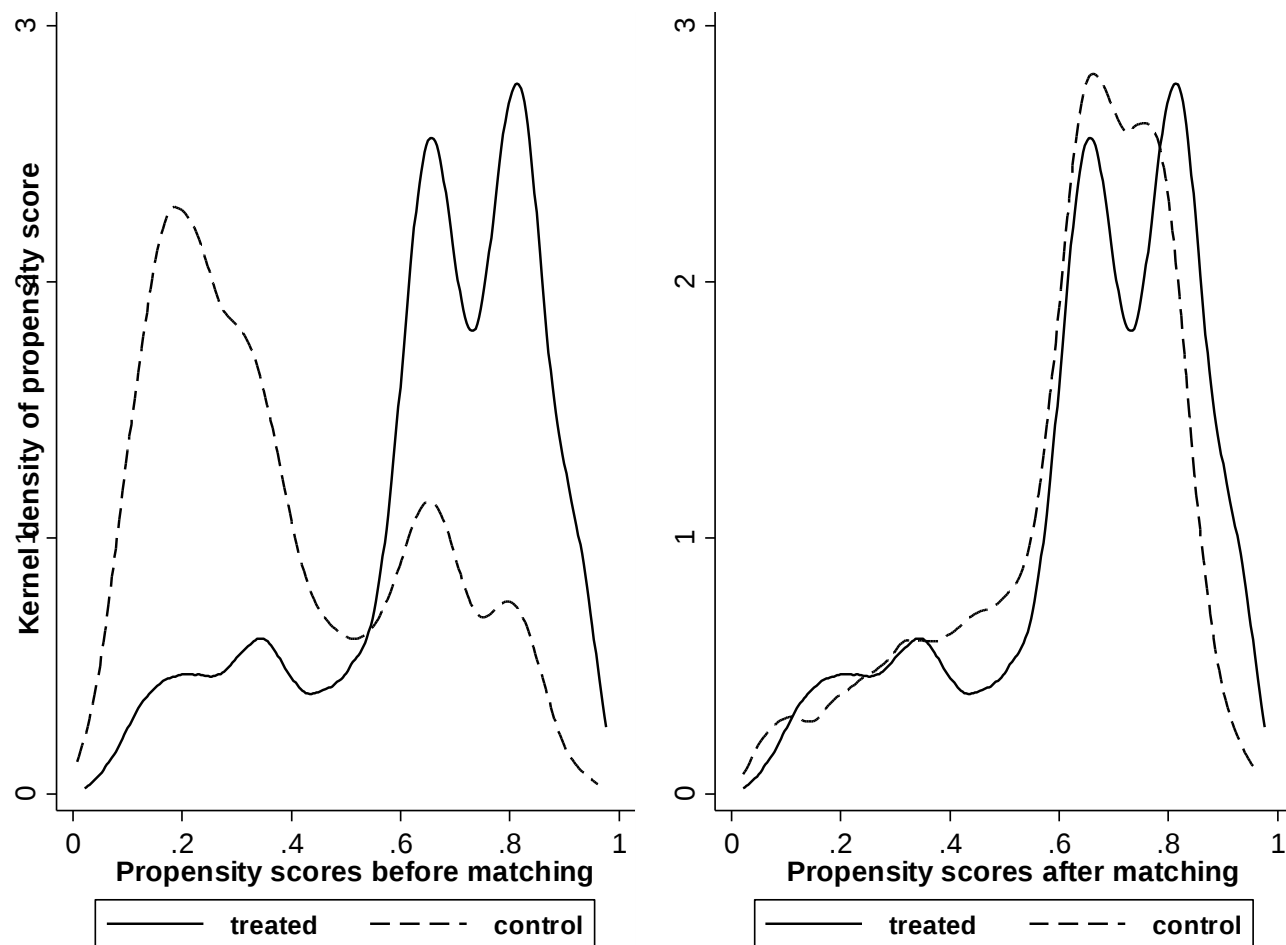


Table A3a. Results of Control Variables in Table 2*

Variable	1	2	3	4	5	6	7	8
Panel A: First-stage results of Heckman (random effects probit) after PSM								
Age	-.002 (.001)	-.001 (.001)	-.001 (.001)	-.002 (.001)	-.001 (.001)	-.001 (.001)	-.002 (.001)	-.002 (.001)
Educational attainment	-.020** (.007)	-.012+ (.007)	-.012 (.007)	-.017* (.007)	-.013+ (.007)	-.014+ (.007)	-.018* (.007)	-.023** (.008)
Current communist ideology	-.007 (.004)	-.004 (.005)	-.004 (.005)	-.004 (.005)	-.005 (.005)	-.004 (.005)	-.005 (.005)	-.005 (.005)
Government work experience	.130** (.034)	.104** (.039)	.113** (.039)	.082* (.040)	.110** (.040)	.121** (.040)	.147** (.040)	.147** (.041)
Foreign experience	.086* (.037)	.101** (.038)	.102** (.038)	.114** (.038)	.102** (.038)	.100** (.038)	.097* (.038)	.110** (.039)
Firm size	.009 (.007)	.005 (.007)	.006 (.007)	.003 (.007)	.005 (.007)	.005 (.007)	.005 (.007)	.003 (.007)
Financial leverage	.114* (.050)	.110* (.051)	.104* (.051)	.116* (.052)	.116* (.052)	.113* (.051)	.108* (.052)	.116* (.053)
Firm age	.005+ (.003)	.011** (.003)	.010** (.003)	.010** (.003)	.012** (.003)	.012** (.003)	.008* (.003)	.009* (.004)
Manufacturing firm	-.023 (.047)	-.014 (.048)	-.016 (.048)	-.017 (.048)	-.019 (.048)	-.010 (.048)	.004 (.048)	-.001 (.049)
Industry average (for corresponding dependent variable)	1.094** (.312)	.160 (.342)	.165 (.342)	.304 (.347)	.154 (.344)	.228 (.342)	.511 (.345)	.588+ (.352)
Poor Internet coverage	-.022 (.076)	-.021 (.077)	-.025 (.077)	-.004 (.078)	-.024 (.078)	-.022 (.077)	-.034 (.078)	-.019 (.080)
Density of CPC membership	.253 (.172)	.258 (.175)	.253 (.175)	.268 (.177)	.224 (.177)	.267 (.176)	.215 (.178)	.194 (.181)
GDP per capita (logged)	.016** (.005)	-.002 (.005)	-.001 (.005)	-.002 (.005)	.002 (.005)	-.001 (.005)	.012* (.005)	.018** (.005)
Population growth	-.487 (.395)	.265 (.417)	.245 (.417)	.046 (.421)	.020 (.422)	.252 (.417)	-.296 (.423)	-.698 (.431)
Institutional development	-.003 (.004)	-.008+ (.004)	-.008+ (.004)	-.008+ (.004)	-.010* (.004)	-.008+ (.004)	-.005 (.004)	-.005 (.005)
R&D investment (0/1)	.019 (.024)	.041+ (.024)	.039 (.024)	.036 (.025)	.037 (.025)	.038 (.024)	.031 (.025)	.019 (.025)
Firm performance (return on assets)	-.017 (.047)	-.217** (.050)	-.212** (.050)	-.051 (.050)	-.195** (.050)	-.223** (.050)	-.168** (.050)	-.013 (.051)
Industry competitiveness	.129 (.092)	.099 (.094)	.112 (.094)	.132 (.096)	.107 (.095)	.101 (.094)	.139 (.096)	.184+ (.098)

Industry growth	-.009 (.018)	-.007 (.017)	-.007 (.017)	-.012 (.018)	-.007 (.017)	-.007 (.017)	-.007 (.018)	-.009 (.018)
Panel B: Second-stage results of Heckman (random effects) after PSM								
Age	.000 ⁺ (.000)	.000 (.000)	.000 (.000)	.000 (.000)	.000 (.000)	.000 (.000)	.000 (.000)	.000 (.000)
Educational attainment	.005 ^{**} (.001)	.003 ^{**} (.001)	.003 ^{**} (.001)	.004 ^{**} (.001)	.003 ^{**} (.001)	.003 ⁺ (.001)	.003 ^{**} (.001)	.002 (.001)
Current communist ideology	.001 (.001)	.001 (.001)	.001 (.001)	.001 (.001)	.001 (.001)	.001 ⁺ (.001)	.001 (.001)	.001 (.001)
Government work experience	-.017 ^{**} (.005)	-.020 ^{**} (.006)	-.017 ^{**} (.006)	-.021 ^{**} (.006)	-.020 ^{**} (.006)	-.016 ^{**} (.006)	-.020 ^{**} (.006)	-.004 (.008)
Foreign experience	-.015 [*] (.006)	-.009 (.006)	-.009 (.006)	-.013 [*] (.006)	-.009 (.006)	-.010 ⁺ (.006)	-.009 (.006)	-.005 (.007)
Firm size	-.000 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)
Financial leverage	-.011 (.009)	.001 (.008)	-.000 (.008)	-.003 (.008)	.001 (.008)	.002 (.008)	.001 (.008)	.005 (.009)
Firm age	-.001 [*] (.000)	-.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)	-.000 (.001)	-.001 (.001)	-.000 (.001)
Manufacturing firm	.008 (.008)	.013 ⁺ (.008)	.012 (.008)	.013 ⁺ (.008)	.013 ⁺ (.008)	.012 (.007)	.013 ⁺ (.008)	.011 (.007)
Industry average (for corresponding dependent variable)	.000 (.013)	-.001 (.012)	-.002 (.012)	-.001 (.012)	-.001 (.012)	-.002 (.012)	-.001 (.012)	-.006 (.012)
Poor Internet coverage	-.041 (.030)	-.023 (.028)	-.024 (.028)	-.030 (.028)	-.024 (.028)	-.015 (.027)	-.023 (.028)	-.012 (.027)
Density of CPC membership	-.002 ^{**} (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)	.000 (.001)
GDP per capita (logged)	.051 (.069)	.002 (.064)	-.006 (.064)	.011 (.064)	.016 (.064)	-.028 (.063)	-.000 (.065)	-.053 (.066)
Population growth	.000 (.001)	-.000 (.001)	-.000 (.001)	.000 (.001)	-.000 (.001)	-.000 (.001)	-.000 (.001)	-.000 (.001)
Institutional development	.188 ^{**} (.069)	-.027 (.065)	-.026 (.065)	-.026 (.065)	-.019 (.065)	-.031 (.064)	-.027 (.065)	-.020 (.063)

⁺ $p < .10$; ^{*} $p < .05$; ^{**} $p < .01$; intercept not reported to save space.

* 18,424 out of 20,564 observations are on the common support and are thus used for PSM analysis; 2,781 out of 3,115 observations are on the common support and are thus used for PSM analysis.

Table A3b. Results of Control Variables in Table 3*

Variable	1	2	3	4	5	6	7	8
Panel A: First-stage results of Heckman (random effects probit) after PSM								
Age	-.001 (.002)	-.001 (.002)	-.001 (.002)	-.001 (.002)	-.001 (.002)	-.001 (.002)	-.001 (.002)	-.001 (.002)
Educational attainment	.001 (.009)	.005 (.009)	.004 (.009)	.004 (.009)	.005 (.009)	.005 (.009)	.004 (.009)	.004 (.009)
Current communist ideology	.000 (.005)	.002 (.006)	.002 (.006)	.002 (.006)	.002 (.006)	.002 (.006)	.001 (.006)	.002 (.006)
Government work experience	-.045 (.046)	-.105* (.053)	-.080 (.054)	-.103+ (.054)	-.097+ (.054)	-.112* (.054)	-.096+ (.054)	-.067 (.054)
Foreign experience	.024 (.041)	.032 (.042)	.032 (.042)	.032 (.042)	.029 (.042)	.032 (.042)	.032 (.042)	.028 (.042)
Firm size	.024** (.008)	.022** (.008)	.022** (.008)	.021* (.008)	.022** (.008)	.021* (.008)	.022** (.008)	.023** (.008)
Financial leverage	-.051 (.062)	-.047 (.063)	-.058 (.063)	-.046 (.063)	-.040 (.063)	-.047 (.063)	-.047 (.063)	-.050 (.063)
Firm age	-.001 (.003)	.003 (.004)	.002 (.004)	.003 (.004)	.004 (.004)	.002 (.004)	.003 (.004)	.003 (.004)
Manufacturing firm	.068 (.055)	.065 (.055)	.062 (.055)	.068 (.055)	.063 (.056)	.065 (.055)	.067 (.055)	.063 (.056)
Industry average (for corresponding dependent variable)	-.189* (.092)	-.179+ (.093)	-.192* (.093)	-.178+ (.093)	-.186* (.093)	-.182* (.093)	-.180+ (.093)	-.203* (.093)
Poor Internet coverage	-.199 (.209)	-.206 (.210)	-.228 (.211)	-.205 (.210)	-.243 (.211)	-.207 (.210)	-.212 (.210)	-.277 (.212)
Density of CPC membership	-.004 (.006)	-.013+ (.007)	-.009 (.007)	-.012+ (.007)	-.008 (.007)	-.013+ (.007)	-.012+ (.007)	-.002 (.007)
GDP per capita (logged)	.301 (.467)	.092 (.481)	.129 (.482)	.069 (.482)	-.096 (.484)	.105 (.482)	.041 (.482)	-.104 (.486)
Population growth	.005 (.006)	-.001 (.006)	-.001 (.006)	-.001 (.006)	-.002 (.006)	-.002 (.006)	-.001 (.006)	-.000 (.006)
Institutional development	-.002 (.028)	-.006 (.028)	-.007 (.028)	-.005 (.028)	-.008 (.028)	-.005 (.028)	-.006 (.028)	-.011 (.028)
R&D investment (0/1)	-.079 (.091)	-.082 (.093)	-.080 (.093)	-.066 (.094)	-.049 (.094)	-.081 (.093)	-.077 (.093)	-.035 (.094)
Firm performance (return on assets)	-.076 (.101)	-.058 (.102)	-.044 (.103)	-.060 (.102)	-.064 (.103)	-.059 (.102)	-.057 (.102)	-.048 (.103)
Industry competitiveness	-.066 (.064)	-.066 (.067)	-.059 (.067)	-.067 (.067)	-.053 (.066)	-.066 (.067)	-.068 (.067)	-.048 (.066)
Industry growth	1.062** (.363)	.434 (.378)	.396 (.380)	.443 (.378)	.351 (.385)	.418 (.378)	.455 (.378)	.332 (.388)

Inward internationalization [†]	3.916** (.146)	3.857** (.148)	3.872** (.148)	3.853** (.148)	3.908** (.149)	3.861** (.148)	3.850** (.148)	3.917** (.149)
Panel B: Second-stage results of Heckman (random effects) after PSM								
Age	.001 (.001)	.000 (.001)	.000 (.001)	.000 (.001)	.000 (.001)	.000 (.001)	.000 (.001)	.000 (.001)
Educational attainment	.006 (.006)	.004 (.006)	.004 (.006)	.004 (.006)	.003 (.006)	.004 (.006)	.004 (.006)	.003 (.006)
Current communist ideology	-.006 (.004)	-.003 (.003)	-.003 (.003)	-.003 (.003)	-.003 (.003)	-.003 (.003)	-.003 (.003)	-.003 (.003)
Government work experience	.091** (.034)	-.018 (.036)	-.014 (.036)	-.018 (.036)	-.001 (.036)	-.017 (.036)	-.018 (.036)	-.001 (.041)
Foreign experience	-.034 (.029)	-.000 (.025)	.001 (.025)	-.001 (.025)	-.005 (.025)	-.001 (.025)	-.000 (.025)	-.005 (.027)
Firm size	-.018** (.006)	-.012* (.005)	-.011* (.005)	-.012* (.005)	-.015** (.005)	-.012* (.005)	-.012* (.005)	-.015+ (.008)
Financial leverage	.055 (.044)	.054 (.039)	.047 (.039)	.055 (.039)	.060 (.039)	.054 (.039)	.054 (.039)	.059 (.043)
Firm age	-.003 (.002)	.002 (.003)	.002 (.003)	.002 (.003)	.001 (.003)	.002 (.003)	.002 (.003)	.001 (.003)
Manufacturing firm	.010 (.038)	.017 (.033)	.019 (.033)	.021 (.033)	.009 (.033)	.016 (.033)	.017 (.033)	.012 (.038)
Industry average (for corresponding dependent variable)	.196** (.066)	.118* (.058)	.099+ (.058)	.118* (.058)	.141* (.059)	.119* (.058)	.118* (.058)	.137 (.084)
Poor Internet coverage	.145 (.146)	-.038 (.129)	-.059 (.129)	-.036 (.129)	-.008 (.129)	-.037 (.129)	-.037 (.129)	-.009 (.150)
Density of CPC membership	.012* (.005)	-.006 (.005)	-.005 (.005)	-.005 (.005)	-.005 (.005)	-.006 (.005)	-.006 (.005)	-.004 (.005)
GDP per capita (logged)	.430 (.316)	.750** (.288)	.776** (.288)	.736* (.288)	.762** (.287)	.749** (.288)	.754** (.288)	.767** (.288)
Population growth	.000 (.004)	.004 (.004)	.004 (.004)	.004 (.004)	.004 (.004)	.004 (.004)	.004 (.004)	.004 (.004)
Institutional development	.186* (.088)	-.081 (.080)	-.090 (.080)	-.081 (.080)	-.090 (.080)	-.081 (.080)	-.081 (.080)	-.093 (.080)
Inward internationalization [†]	.820** (.262)	.185 (.255)	.459 (.279)	.220 (.257)	.298 (.338)	.173 (.258)	.176 (.259)	.218 (.151)

+ $p < .10$; * $p < .05$; ** $p < .01$; intercept not reported to save space.

* 12,643/2,171 out of 14,119/2,476 observations are on the common support and are thus used for PSM analysis in the first/second stage.

† Foreign investment over total assets was instrumented by presence of foreign concession, whose test is based on joint statistical independence (Zhang, Marquis, and Qiao, 2016).

Test results suggested that the instruments are highly correlated with foreign investments and satisfy the exogeneity and exclusion restriction.

Appendix B: Robustness Checks and Additional Analyses

Table B1a. Results from Heckman Model: Inward Internationalization (1993–2012)*

Variable	1	2	3	4	5	6	7
Panel A: First-stage results of Heckman (random effects probit); controls are the same as Panel A of table A3a and not reported							
Ideological imprint (H1)	-.505** (.032)	-.589** (.038)	-1.198** (.051)	-.431** (.066)	-2.125** (.230)	-4.822** (.559)	-3.907** (.303)
Political involvement	.102** (.027)	.059* (.029)	.096** (.027)	.107** (.027)	.099** (.027)	.086** (.027)	.051+ (.029)
Government appropriation	1.149** (.109)	1.137** (.109)	.520** (.151)	1.003** (.111)	1.131** (.110)	.857** (.113)	.609** (.152)
Social network	.146** (.026)	.147** (.026)	.116** (.026)	.029 (.027)	.145** (.026)	.120** (.026)	.012 (.027)
Post 2001	.041 (.044)	.037 (.044)	.001 (.044)	.023 (.044)	.016 (.045)	.080+ (.044)	.011 (.046)
Regional FDI intensity	.503** (.047)	.498** (.047)	.402** (.048)	.453** (.048)	.493** (.047)	.115* (.057)	.173** (.058)
Ideological imprint × Political involvement (H2a)		.280** (.066)					.308** (.073)
Ideological imprint × Government appropriation (H2b)			4.521** (.238)				3.897** (.262)
Ideological imprint × Social network (H3)				.186** (.023)			.170** (.029)
Ideological imprint × Post 2001 (H4a)					.440** (.073)		.102* (.024)
Ideological imprint × Regional FDI intensity (H4b)						1.046** (.102)	.311** (.110)
Number of observations	20,564	20,564	20,564	20,564	20,564	20,564	20,564
Goodness of fit (pseudo R ²)	.367	.397	.397	.398	.401	.404	.423
Panel B: Second-stage results of Heckman (random effects linear); controls are the same as Panel B of table A3a and not reported							
Inverse Mills ratio	-.074** (.008)	-.070** (.008)	-.104** (.011)	-.069** (.008)	-.062** (.008)	-.072** (.012)	.003 (.045)
Ideological imprint (H1)	-.128** (.007)	-.111** (.008)	-.160** (.010)	-.181** (.069)	-.216** (.053)	-.126** (.013)	-.516** (.140)
Political involvement	.031** (.004)	.023** (.005)	.028** (.004)	.031** (.004)	.032** (.004)	.031** (.004)	.026** (.005)
Government appropriation	.021 (.017)	.024 (.017)	.092** (.024)	.024 (.017)	.025 (.017)	.022 (.018)	.034 (.033)
Social network	.006 (.004)	.007 (.004)	.006 (.004)	.006 (.004)	.006 (.004)	.006 (.004)	.008* (.004)
Post 2001	.018** (.006)	.016** (.006)	.021** (.006)	.018** (.006)	.016** (.006)	.018** (.006)	.018** (.006)
Regional FDI intensity	.028** (.008)	.028** (.008)	.018* (.008)	.030** (.008)	.029** (.007)	.027** (.009)	.016 (.010)
Ideological imprint × Political involvement (H2a)		.490** (.110)					.651** (.113)
Ideological imprint × Government appropriation (H2b)			.186** (.042)				.978** (.118)
Ideological imprint × Social network (H3)				.311** (.069)			.236* (.093)
Ideological imprint × Post 2001 (H4a)					.136** (.014)		.151** (.015)
Ideological imprint × Regional FDI intensity (H4b)						.004+ (.002)	.100+ (.061)
Number of observations	3,115	3,115	3,115	3,115	3,115	3,115	3,115
Goodness of fit (between R ²)	.325	.328	.330	.330	.348	.325	.361

+ $p < .10$; * $p < .05$; ** $p < .01$; intercept not reported to save space.

* The results are from the original sample with unique identifiers (20,564 firm-year observations) and are similar to what were reported with the PSM approach in tables 2 and 3 and tables A3a and A3b.

Table B1b. Results from Heckman Model: Outward Internationalization (1993–2012)*

Variable	1	2	3	4	5	6	7
Panel A: First-stage results of Heckman (random effects probit); controls are the same as Panel A of table A3b and not reported							
Ideological imprint (H1)	-.317** (.035)	-.448** (.042)	-.397** (.059)	-1.118** (.093)	-1.869** (.269)	-1.390** (.351)	-1.247** (.127)
Political involvement	.047 (.033)	-.039 (.036)	.046 (.033)	.058+ (.033)	.048 (.033)	.045 (.033)	-.025 (.036)
Government appropriation	.143 (.198)	.089 (.199)	-.282 (.325)	-.106 (.200)	.163 (.198)	.041 (.205)	-.334 (.325)
Social network	.218** (.032)	.222** (.032)	.216** (.032)	.090** (.033)	.218** (.032)	.216** (.032)	.093** (.034)
Post 2001	.042 (.049)	.037 (.049)	.044 (.049)	.038 (.049)	.068 (.050)	.050 (.049)	.073 (.051)
Regional FDI intensity	.165** (.058)	.154** (.058)	.151** (.058)	.124* (.058)	.162** (.058)	.102 (.066)	.016 (.067)
Ideological imprint × Political involvement (H2a)		.458** (.074)					.564** (.077)
Ideological imprint × Government appropriation (H2b)			.691+ (.409)				1.744** (.603)
Ideological imprint × Social network (H3)				.127** (.031)			.154** (.035)
Ideological imprint × Post 2001 (H4a)					.164* (.077)		.088* (.042)
Ideological imprint × Regional FDI intensity (H4b)						.215* (.107)	.091** (.032)
Number of observations	14,119	14,119	14,119	14,119	14,119	14,119	14,119
Goodness of fit (pseudo R ²)	.450	.475	.474	.477	.484	.486	.499
Panel B: Second-stage results of Heckman (random effects linear); controls are the same as Panel B of Table A3b and not reported							
Inverse Mills ratio	-.391** (.086)	-.303** (.095)	-.378** (.086)	-.545** (.114)	-.398** (.087)	-.390** (.087)	-.577 (.456)
Ideological imprint (H1)	-.623** (.026)	-.674** (.035)	-.657** (.040)	-.891** (.118)	-.639** (.045)	-.625** (.036)	-1.415** (.478)
Political involvement	.172** (.020)	.153** (.022)	.172** (.020)	.161** (.021)	.171** (.020)	.172** (.020)	.158** (.024)
Government appropriation	.540** (.123)	.538** (.123)	.356+ (.209)	.563** (.123)	.536** (.123)	.537** (.132)	.417+ (.242)
Social network	-.041+ (.021)	-.028 (.022)	-.040+ (.021)	-.046* (.021)	-.042+ (.021)	-.041+ (.021)	-.048 (.039)
Post 2001	.012 (.031)	.012 (.031)	.013 (.031)	.005 (.031)	.009 (.032)	.012 (.031)	-.005 (.041)
Regional FDI intensity	.032 (.037)	.039 (.037)	.027 (.037)	.019 (.037)	.032 (.037)	.031 (.041)	.028 (.041)
Ideological imprint × Political involvement (H2a)		.313** (.052)					.510** (.072)
Ideological imprint × Government appropriation (H2b)			.280+ (.158)				.787** (.269)
Ideological imprint × Social network (H3)				.234* (.115)			.062 (.040)
Ideological imprint × Post 2001 (H4a)					.021+ (.011)		.014+ (.008)
Ideological imprint × Regional FDI intensity (H4b)						.004* (.002)	.007+ (.004)
Number of observations	2,476	2,476	2,476	2,476	2,476	2,476	2,476
Goodness of fit (between R ²)	.366	.367	.366	.366	.366	.366	.367

+ $p < .10$; * $p < .05$; ** $p < .01$; intercept not reported to save space.

* The results are from the original sample with unique identifiers (20,564 firm-year observations) and are similar to what were reported with the PSM approach in tables 2 and 3 and tables A3a and A3b.

A Counterfactual Analysis for Understanding Effects of Profitability from Internationalization (1993–2012)

To better triangulate the imprinting processes we theorize, we also examined whether entrepreneurs with a communist ideological imprint tend to ignore profitable business opportunities from internationalization, showing whether the imprint dominates self-interest in considering foreign cooperation, i.e., whether the information filter by the communist ideological imprint motivates cognition. We measured profitable internationalization opportunities by calculating profitability differentials between internationalized firms and their non-internationalized/domestic counterparts of focal firms' institutional equivalents, i.e., those in the same geographical location (province) and industry as their counterfactuals, respectively (Marquis and Tilcsik, 2016). Then we interacted these two proxies of information of profitable internationalization with the communist ideological imprint to predict the tendency (hazard rate) of internationalization of the focal firm with a Cox proportional hazard model. To estimate the Cox model, we stratified the model at the industry–province level to alleviate unobservable heterogeneity concerns and employed the graphic method (Andersen, 1982) and partial residuals approach (Schoenfeld, 1982) to ensure that the proportionality assumption holds. We also clustered standard errors at the industry–province level to obtain the most robust results (Lin and Wei, 1989).

Table B2 shows that profit premiums of the same industry, same province, and their interaction are all positively associated with firms' tendency to internationalize. But the interaction terms of all three variables with ideological imprint are negative and significant ($p < .05$), suggesting that firm leaders filtered information about profitable opportunities via internationalization with a communist ideological imprint. These results lend support to our theorizing that the communist ideological imprint acts as an information filter that motivates cognition—entrepreneurs eschew cooperation with foreign capitalists even when it is in their economic self-interest. Results from the Heckman model are similar.

Table B2. Predicting Inward and Outward Internationalization with Counterfactual Profits (1993–2012)*

Analysis Variable	Inward Internationalization				Outward Internationalization			
	1	2	3	4	5	6	7	8
	Cox model		Heckman 1st and 2nd stages		Cox model		Heckman 1st and 2nd stages	
	Single event	Repeated events	RE probit [†]	RE linear	Single event	Repeated events	RE probit	RE linear
Ideological imprint	–.206** (.073)	–.444** (.122)	–.250** (.052)	–.035** (.008)	–.167* (.075)	–.767* (.340)	–3.892** (1.126)	–.513** (.138)
Profit premium by geography	2.443** (.288)	.024* (.011)	.024** (.006)	.003** (.001)	2.265** (.306)	.086* (.036)	1.163+ (.677)	.211* (.094)
Profit premium by industry	1.842** (.269)	.022** (.008)	.027** (.008)	.006** (.001)	1.415** (.286)	.086* (.034)	1.067+ (.579)	.306** (.092)
Profit premium by geography × Profit premium by industry	3.834** (.404)	.157** (.058)	.074** (.015)	.008** (.001)	2.535** (.424)	.155** (.054)	1.213** (.305)	.174+ (.100)
Ideological imprint × Profit premium by geography	–4.493** (.352)	–.221** (.070)	–.691** (.163)	–.015** (.005)	–3.707** (.360)	–.083** (.016)	–1.515** (.477)	–1.544** (.282)
Ideological imprint × Profit premium by industry	–4.695** (.348)	–1.735** (.357)	–.326** (.097)	–.013* (.006)	–3.841** (.360)	–1.868** (.643)	–2.029** (.333)	–1.055* (.506)
Ideological imprint × profit premium by geography × Profit premium by industry	–4.456** (.364)	–4.291** (.567)	–.772* (.385)	–.118* (.054)	–3.904** (.376)	–3.034** (.590)	–3.608** (.561)	–1.220** (.278)
Number of observations	15,681	20,564	20,564	3,115	10,217	14,119	14,119	2,476
Goodness of fit (pseudo R ²)	.125	.186	.330		.104	.193	.494	
Goodness of fit (between R ²)				.303				.421

+ $p < .10$; * $p < .05$; ** $p < .01$; intercept not reported to save space. The sample size for the Cox model single event history analysis varies because the right-censored firms are dropped whenever they internationalized their firms.

* The repeated events Cox proportional hazard model retains all observations even if the focal firm/individual has already adopted the action, and therefore numbers of observations are the same as tables A3a and A3b for corresponding dependent variables. Standard errors are clustered at the industry–province level to obtain the most robust results (Lin and Wei, 1989). Controls and moderators are the same as tables A3a and A3b, respectively (not reported); interaction terms are excluded.

† “RE” refers to random effects.

Diff-in-diff and Related Estimation of Internationalization

We provide a finer-grained analysis of communist ideological imprint and also rule out cohort effects as an alternative explanation—it is not only those born before 1978 who are antagonistic to foreign cooperation but it is the importance of the socialization of joining the CPC—based on a diff-in-diff analysis. We distinguished those born before and after 1978, in addition to communist ideology. Therefore we analyzed four groups of entrepreneurs: pre- and post-1978 communists and non-communists. Unlike traditional diff-in-diff estimation, our diff-in-diff analysis is not about the event of imprinting but about whether the entrepreneur had the ideological imprint (treatment) or not (control), and the “event” was whether opening up in 1978 invalidates the negative impact of current ideology on internationalization such that the post-1978 (reformist) government encourages capital and foreign cooperation more fully. Our approach resembles the diff-in-diff estimation for repeated cross sections, which utilizes between-differences of individuals rather than within ones, i.e., different individuals before and after the events (see Abadie, 2005: 9, for summary of exemplar studies).

We found that pre-1978 communists are least prone to internationalization; we present results in Panel A of table B3. We also combined the diff-in-diff analysis with PSM as going through the imprinting process might be self-selected and thus endogeneity could be an issue. The PSM approach helps generate a random sample in terms of entrepreneurs with or without a communist ideological imprint based on observable variables. The results are shown in Panel B of table B3 and are consistent with the main analyses. Finally, we distinguished entrepreneurs’ Party age (years since they joined the CPC) to gauge when they received the communist ideological imprinting process. We had a set of three dummy variables: joined the CPC before 1978, joined the CPC after 1978 and before they founded their ventures, and joined the CPC after they founded their ventures (an alternative proxy for current ideology). The results are shown in Panel C of table B3, suggesting that the ideological imprint formed before 1978 has the strongest negative effect on internationalization, while that formed after 1978 but before entrepreneurs founded their ventures is much weaker. Contemporary ideology does not make a difference statistically or economically.

Table B3. Diff-in-diff and Related Estimation of Internationalization*

Panel A: Diff-in-diff estimation of internationalization				
Variable	Inward internationalization		Outward internationalization	
	0/1 1	Value 2	0/1 3	Value 4
CPC membership	-.082 (.063)	-.001 (.002)	-1.075 (.824)	-.006 (.008)
Born before 1978	-.159* (.062)	-.009* (.003)	-2.704* (1.261)	-.071* (.032)
CPC membership × Born before 1978	-.234** (.042)	-.021** (.006)	-4.858** (.624)	-.277** (.069)
Inverse Mills ratio [†]		-1.502** (.441)		-2.315** (.447)
Number of observations	20,564	3,115	20,564	2,476
Goodness of fit (pseudo R ²)	.399		.476	
Goodness of fit (between R ²)		.315		.404
Panel B: Diff-in-diff estimation of internationalization combined with propensity score matching				
Variable	Inward internationalization		Outward internationalization	
	Binary	Continuous	Binary	Continuous
CPC membership	-.005 (.008)	-.001 (.004)	-.894 (.837)	-.012 (.013)
Born before 1978	-.035* (.016)	-.005* (.002)	-1.524* (.655)	-.041* (.020)
CPC membership × Born before 1978	-.475** (.074)	-.015** (.003)	-5.910** (1.297)	-.380** (.113)
Inverse Mills ratio [†]		-1.468** (.390)		-2.417** (.443)
Number of observations	18,424	2,781	12,643	2,171
Goodness of fit (pseudo R ²)	.370		.372	
		.281		.286
Panel C: A set of dummy variables to measure ideological imprint and current ideology				
Variable	Inward internationalization		Outward internationalization	
	Binary	Continuous	Binary	Continuous
Joined CPC before 1978	-.466** (.144)	-.043** (.008)	-3.863** (1.148)	-.347** (.107)
Joined CPC after 1978 but before they founded their venture	-.025** (.005)	-.002* (.001)	-.046** (.016)	-.031* (.015)
Joined CPC after founding their venture (current political ideology)	-.000 (.007)	-.000 (.003)	-.010 (.043)	-.000 (.004)
Inverse Mills ratio [‡]		-1.376** (.318)		-2.861** (1.050)
Number of observations	20,564	3,115	14,119	2,476
Goodness of fit (pseudo R ²)	.364		.442	
Goodness of fit (between R ²)		.271		.386

[†] $p < .10$; * $p < .05$; ** $p < .01$; intercept not reported to save space.
* Numbers of observations for the PSM results in Panel B vary because of the matching. Controls and moderators are the same as tables A3a and A3b, respectively (not reported); interaction terms are excluded.
[†] Inverse Mills ratio was from the first stage including two additional variables—*born before 1978* and *CPC membership × Born before 1978* corresponding to the diff-in-diff setting.
[‡] Inverse Mills ratio was from the first stage including the corresponding three binary variables in the dummy approach.

Other Endogeneity Concerns and Moderators Unrelated to Information Filtering

We tried a number of additional approaches to address a variety of endogeneity concerns. For instance, we use the variation of parents' government work experience of entrepreneurs, which highly correlates with CPC membership (Bian, Shu, and Logan, 2001), as an instrumental variable. Parents' CPC government work experience is exogenous and not affected by entrepreneurs' later new ventures. Studies have shown intergenerational transmissions of ideology (Jost, Federico, and Napier, 2009), and CPC membership is required for governmental work (Wu and Treiman, 2007). Lastly, parents' work experience is at best weakly correlated with firms' internationalization, given that nearly all entrepreneurs' parents—assuming they gave birth to their children in their twenties—are in their late eighties. Therefore the instrumental variable is valid conceptually. We still ran a first-stage F-test, test for the exclusion restriction, according to Zhang, Marquis, and Qiao (2016), and obtained supporting evidence that our instrumental variable is valid. Results are reported in columns 1 and 2 of table B4 and are similar to our main results.

Endogeneity issues may also arise because of omitted control variables. Therefore we also controlled for (1) state-owned-enterprise work experience and political rank to indicate whether the focal entrepreneur ever worked as a government official, and (2) the exclusion restrictions, i.e., technological resource, firm performance, industry competitiveness, and growth (Villalonga and McGahan, 2005). We report results in columns 3 to 6 of table B4, which are similar in terms of sign, significance, and magnitude on the communist ideological imprint. Additional control variables are insignificant both respectively and jointly, however, and a Hausman test suggested that these variables did not change our results in terms of magnitude or significance (Hausman, 1978).

In addition, we interacted firm age, entrepreneurs' age, and time trend with their ideological imprint. If the interaction terms are significant and negative, then our arguments that the availability and credibility of contradictory information are two key conditions for imprint decay would be undermined; the imprint decays automatically over time. The last six models (models 7 to 12) of both panels in table B4 suggest this was not the case; all three interaction terms are insignificant, and the effect size is less than .0001, or less than 5 percent of other moderating effects. Therefore the imprinting effect does not decay with an increase in firm age, entrepreneurs' age, or time.

Table B4. Results from Heckman Model of Internationalization (1993-2012): Other Endogeneity Concerns, Additional Control Variables, and Interactive Effects*

Approach Variable	1 (B) [†] Instrumental variable Parent is CPC member	2 (C)	3 (B) Additional control variables Entrepreneur [‡]	4 (C) Firm/industry	5 (B)	6 (C)	7 (B)	8 (C)	9 (B)	10 (C)	11 (B)	12 (C)
Panel A: Dependent variable—foreign investment over total investment Controls and other moderators are the same as table A3a and not reported												
Ideological imprint (H1)	-.250** (.065)	-.030** (.010)	-.271** (.100)	-.035** (.007)	/	-.022* (.011)	-.278** (.068)	-.026** (.009)	-.241** (.067)	-.030** (.005)	-.337** (.108)	-.022** (.007)
State-owned-enterprise work experience			-.001 (.003)	-.001 (.001)								
Political rank when working as a government official			-.000 (.001)	-.012 (.009)								
Tech resources					/§	.002 (.006)						
Firm performance					/	.01 (.007)						
Industry competitiveness					/	.001 (.003)						
Industry growth					/	-.002 (.004)						
Firm age							.000 (.001)	.000 (.000)				
Ideological imprint × Firm age							.000 (.003)	-.000 (.001)				
Entrepreneur's age									-.002 (.001)	.000 (.000)		
Ideological imprint × Entrepreneur's age									.000 (.001)	.000 (.000)		
Time trend											.002 (.011)	-.001 (.004)
Ideological imprint × Time trend											-.000 (.002)	-.000 (.000)
Weak instrument test	32.394 (passed)											
Exclusion restriction test <i>p</i> value	.192											
Number of observations	3,290	494	3,701	666	/	3,115	20,564	3,115	20,564	3,115	20,564	3,115
Goodness of fit (pseudo R ²)	.405		.376		/		.368		.369		.369	
Goodness of fit (between R ²)		.323		.280		.313		.316		.318		.312

Panel B: Dependent variable—Overseas assets over total assets												
Controls and other moderators are the same as table A3b and not reported												
Ideological imprint (H1)	−3.776*	−.259**	−2.342**	−.212**	/	−.277**	−3.231**	−.364**	−2.952**	−.267*	−3.501**	−.290**
	(1.675)	(.086)	(.415)	(.059)	/	(.069)	(.759)	(.119)	(.979)	(.106)	(1.126)	(.096)
State-owned enterprise work experience			−.000	−.006								
			(.010)	(.042)								
Political rank when working as a government official			−.005	−.001								
			(.009)	(.005)								
Tech resources					/	.152						
					/	(.096)						
Firm performance					/	.043						
					/	(.108)						
Industry competitiveness					/	.000						
					/	(.093)						
Industry growth					/	−.000						
					/	(.001)						
Firm age							−.012	−.011				
							(.008)	(.010)				
Ideological imprint × Firm age							.000	.000				
							(.004)	(.001)				
Entrepreneur's age									−.002	.007		
									(.002)	(.006)		
Ideological imprint × Entrepreneur's age									.000	−.000		
									(.005)	(.002)		
Time trend											−.004	−.002
											(.007)	(.003)
Ideological imprint × Time trend											−.000	.000
											(.004)	(.007)
Weak instrument test	30.628 (passed)											
Exclusion restriction test <i>p</i> value	.219											
Number of observations	2,259	407	2,542	457	/	2,476	14,119	2,476	14,119	2,476	14,119	2,476
Goodness of fit (pseudo R ²)	.375		.384		/		.453		.453		.452	
Goodness of fit (between R ²)		.368		.365		.382		.383		.383		.383

* $p < .10$; * $p < .05$; ** $p < .01$; intercept not reported to save space.

* The number of observations changes because parental information was available only in 1993, 1995, and 1997. Therefore 3,290 observations were used to investigate the binary inward internationalization variable, and 494 observations were used to investigate the continuous inward internationalization variable; 2,259 observations were used to investigate the binary outward internationalization variable, and 407 observations were used to investigate the continuous outward internationalization variable.

† B: binary measure of dependent variable, C: continuous measure of dependent variable.

‡ The two additional control variables of entrepreneurs did not exist until Year 2004. Therefore 3,701 observations were used to investigate the binary inward internationalization variable, and 666 observations were used to investigate the continuous inward internationalization variable; 2,542 observations were used to investigate the binary outward internationalization variable, and 457 observations were used to investigate the continuous outward internationalization variable.

§ These “additional controls” are exclusion restrictions, and thus the regression results are omitted to avoid duplication with results of model 2 in tables A3a and A3b.

Effects of Heterogeneity of Communist Ideological Imprint on Internationalization

The communist ideological imprint may exhibit some heterogeneity, the substantial magnitude of which may threaten the validity of our moderators as decaying factors of the communist ideological imprint. To address this issue, we considered regional variation—coastal region, geographic proximity to special economic zones, and local communist density during the imprinting period (Liu, Buck, and Shu, 2005; Luo, Xue, and Han, 2010; Kung and Chen, 2011)—that may also influence the intensity of the imprinting effect. Likewise, age groups of entrepreneurs may also affect the intensity of communist ideological indoctrination. During the Cultural Revolution (1966–1976), communist ideology was strengthened (Lu, 2004). Therefore the group born and brought up during the Cultural Revolution might exhibit a stronger imprinting effect and thus more heightened information filtering.

The results are reported in table B5, showing that most of these effects do not change the magnitude or significance of the communist ideological imprint. The first three contingencies are insignificant, whereas in some cases birth in the Cultural Revolution period and from the onset of the Cultural Revolution to the year Nixon visited China are marginally significant at the 10-percent level, though the moderating effects are less than 5 percent of other hypothesized moderating effects and hence are not economically significant. The results suggest that the heterogeneity of the ideological imprint does not pose a substantive threat to our analysis.

Table B5. Results from Heckman Model of Internationalization Considering Heterogeneity of Communist Ideological Imprint (1993–2012)

Model	1 (B)*	2 (C)	3 (B)	4 (C)	5 (B)	6 (C)	7 (B)	8 (C)	9 (B)	10 (C)
Contingency	Coastal region		Geographic distance to special economic zone		CPC density during the imprinting period		Cultural Revolution (1966–1976)		Cultural Revolution until Nixon visited China (1966–1972)	
Panel A: Dependent variable—inward internationalization										
Controls and other moderators are the same as table A3a and not reported										
Ideological imprint (H1)	–.168** (.038)	–.031** (.007)	–.161** (.034)	–.030** (.008)	–.150* (.064)	–.024** (.008)	–.194 (.039)	–.028** (.007)	–.234** (.084)	–.023** (.008)
Contingency	–.007* (.003)	–.001* (.000)	–.008** (.003)	–.004 (.003)	–.008* (.003)	–.001** (.000)	–.013** (.004)	–.001+ (.001)	–.008* (.003)	–.000* (.000)
Ideological imprint × Contingency	–.001 (.003)	–.001 (.002)	.000 (.002)	.001 (.003)	–.001 (.002)	–.000 (.003)	–.001 (.001)	–.000 (.001)	–.001 (.001)	–.001+ (.001)
Number of observations	20,564	3,115	20,564	3,115	20,564	3,115	20,564	3,115	20,564	3,115
Goodness of fit (pseudo R ²)	.383		.387		.388		.393		.388	
Goodness of fit (between R ²)		.334		.335		.334		.341		.321
Panel B: Dependent variable—outward internationalization										
Controls and other moderators are the same as table A3b and not reported										
Ideological imprint (H1)	–2.364** (.553)	–.215** (.046)	–2.301** (.553)	–.119** (.065)	–2.573** (.557)	–.162** (.047)	–1.953** (.779)	–.189* (.080)	–1.506** (.353)	–.192** (.072)
Contingency	–.031** (.007)	–.042 (.084)	–.027** (.008)	–.009 (.015)	–.014 (.009)	–.001** (.000)	–.048 (.055)	–.013 (.016)	–.019+ (.011)	–.000 (.000)
Ideological imprint × Contingency	–.000 (.002)	–.000 (.001)	–.001 (.002)	.001 (.002)	.002 (.001)	.003 (.005)	–.001 (.001)	–.001+ (.001)	–.001 (.002)	–.001 (.002)
Number of observations	14,119	2,476	14,119	2,476	14,119	2,476	14,119	2,476	14,119	2,476
Goodness of fit (pseudo R ²)	.450		.483		.453		.467		.460	
Goodness of fit (between R ²)		.403		.380		.375		.408		.397

+ $p < .10$; * $p < .05$; ** $p < .01$; intercept not reported to save space.

* B: binary measure of dependent variable, C: continuous measure of dependent variable.

Predicting Current Political Ideology

An important issue to examine is the extent to which the ideological imprint affects the current political ideology of entrepreneurs. If our model correctly predicts the effect of the ideological imprint and its evolution, then we should expect that (1) the ideological imprint positively affects entrepreneurs' current political ideology because of its persistent influence (Marquis and Tilcsik, 2013), and (2) direct interactions with the contemporary reformer-led government, government-created industry social network involvement, and observing governmental support of internationalization should all negatively moderate this relationship. We report results in table B6 with the same model specification as in Panel A of table 2 (shown in table A3a) and obtained supporting evidence for all of our predictions.

Table B6. Predicting Current Political Ideology: Random Effects Model (1993–2012)*

Variable	1	2	3	4	5	6	7
Ideological imprint	.747** (.224)	.788** (.224)	1.014** (.254)	.728** (.223)	.797** (.224)	.788** (.224)	.818** (.257)
Political involvement	-.585** (.086)	-.521** (.084)	-.487** (.085)	-.483** (.084)	-.374** (.108)	-.503** (.085)	-.325* (.134)
Government appropriation	-1.185** (.045)	-1.256** (.057)	-1.207** (.045)	-1.155** (.045)	-1.202** (.045)	-1.206** (.045)	-1.194** (.058)
Social network	.004 (.016)	.005 (.016)	.006 (.016)	.082** (.018)	.007 (.016)	.006 (.016)	.081** (.018)
Post 2001	.187** (.042)	.270** (.048)	.216** (.042)	.503** (.083)	.388** (.085)	.273** (.052)	.019 (.343)
Regional FDI intensity	-.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)	-.000 (.001)	-.001 (.001)
Ideological imprint × Political involvement		-1.103** (.276)					-1.085** (.277)
Ideological imprint × Government appropriation			-2.848** (.447)				-2.504** (.473)
Ideological imprint × Social network				-.795** (.071)			-.785** (.071)
Ideological imprint × Post 2001					-.197* (.096)		-.182* (.106)
Ideological imprint × Regional FDI intensity						-.161** (.042)	-.129* (.052)
Number of observations	20,564	20,564	20,564	20,564	20,564	20,564	20,564
Goodness of fit (between R ²)	.206	.214	.215	.217	.215	.215	.240

+ $p < .10$; * $p < .05$; ** $p < .01$; intercept not reported to save space.

* Controls are the same as table A3a and are not reported.

Alternative Measures of Key Constructs

We also used alternative measures for some of our proposed constructs. For outward internationalization, we investigated the number of foreign subsidiaries with zero-inflated negative binomial regression, as the dependent variable is overdispersed with more than 10% zeroes (Hausman, Hall, and Griliches, 1984; Blevins, Tsang, and Spain, 2015). Since the Vuong's alpha was positive and significant, the zero-inflated negative binomial model was the most appropriate. We followed suggestions by Blevins, Tsang, and Spain (2015) in analyzing the model. We examined All-China Federation of Industry and Commerce membership as an alternative to the government-created industry social network (Jia, 2014). Similar results as those reported in table 3 can be found in table B7.

Table B7. Results from Zero-inflated Negative Binomial Estimation of Outward Internationalization (2000–2012): Number of Foreign Subsidiaries as the Dependent Variable*

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Ideological imprint (H1)	–.383* (.151)	–.570** (.193)	–.496** (.153)	–.128** (.036)	–.429* (.190)	–1.905* (.802)	–6.548** (1.463)	–.456** (.158)	–.484* (.199)	–.598** (.160)	–.801** (.245)	–.424* (.199)	–.326** (.072)	–6.720** (1.740)
Political involvement	.734* (.327)	.713* (.326)	.725* (.327)	.731* (.327)	.720* (.326)	.734* (.327)	.697* (.326)	.645+ (.334)	.628+ (.333)	.636+ (.334)	.639+ (.334)	.627+ (.333)	.645+ (.334)	.610+ (.333)
Government appropriation	5.311** (.973)	5.356** (.976)	4.437* (1.982)	5.249** (.974)	5.316** (.975)	5.311** (.973)	4.848* (1.954)	5.581** (1.078)	5.602** (1.077)	4.700* (2.079)	5.491** (1.081)	5.596** (1.086)	5.581** (1.078)	5.400* (2.122)
Social network (by APE)	.270** (.040)	.273** (.040)	.270** (.040)	.272** (.040)	.269** (.040)	.270** (.040)	.272** (.040)							
Social network (by ACPIC)								.301** (.038)	.303** (.038)	.301** (.038)	.303** (.038)	.300** (.038)	.301** (.038)	.302** (.038)
Post 2001	1.541** (.327)	1.529** (.327)	1.541** (.327)	1.570** (.328)	1.428** (.327)	1.541** (.327)	1.440** (.329)	1.371** (.340)	1.363** (.339)	1.372** (.340)	1.409** (.342)	1.248** (.340)	1.372** (.340)	1.268** (.342)
Regional FDI intensity	.020** (.004)	.020** (.004)	.020** (.004)	.020** (.004)	.038** (.008)	.020** (.004)	.036** (.008)	.016** (.005)	.016** (.005)	.016** (.005)	.016** (.005)	.034** (.008)	.016** (.005)	.032** (.008)
Ideological imprint × Political involvement (H2a)		.553* (.256)					.534* (.259)		.513* (.240)					.488* (.217)
Ideological imprint × Government appropriation (H2b)			5.204* (2.231)				5.605* (2.215)			5.248* (2.388)				5.175* (2.427)
Ideological imprint × Social network (by APE, H3)				.279* (.125)			.278* (.125)							
Ideological imprint × Social network (by ACPIC, H3)											.336** (.113)			.329** (.115)
Ideological imprint × Post 2001 (H4a)					.028** (.008)		.025** (.008)					.027** (.008)		.024** (.008)
Ideological imprint × Regional FDI intensity (H4b)						.099** (.015)	.095* (.046)						.110** (.037)	.087* (.040)
Number of observations	14,119	14,119	14,119	14,119	14,119	14,119	14,119	14,119	14,119	14,119	14,119	14,119	14,119	14,119
Goodness of fit (pseudo R ²)	.167	.169	.168	.168	.171	.168	.176	.165	.167	.167	.166	.170	.165	.177

+ $p < .10$; * $p < .05$; ** $p < .01$; intercept not reported to save space.

* Controls are the same as table A3a and are not reported.

Robustness across Industries

We considered industry effects by resorting to more-nuanced classifications. Some industries—such as manufacturing—are much more prone to internationalize than are others, such as services and cultural goods. We thus split the sample into manufacturing firms, firms in high-tech or trade-related industries that have an international focus, and firms in other industries, including services, advertisement, consulting, art, retail, etc. We estimated the subsamples under the seemingly unrelated regressions framework (Zellner, 1962) to facilitate cross-equation comparison of effect sizes and report results in table B8. The results are still similar to what were reported in tables 2 and 3. The negative impact of the communist ideological imprint is the strongest in services and cultural goods industries. Meanwhile, we considered time-varying industry variables, such as dynamism, as an additional control variable and obtained supporting results.

Table B8. Subsample Results by Industry from Heckman Model of Internationalization (1993–2012)*

Sample	1 (B) [†]	2 (C)	3 (B)	4 (C)	5 (B)	6 (C)	7 (B)	8 (C)	9 (B)	10 (C)	11 (B)	12 (C)
	Manufacturing		High technology and trade		Others [‡]		Manufacturing		High technology and trade		Others	
Dependent variable	Inward internationalization						Outward internationalization					
	Controls are the same as table A3a and not reported						Controls are the same as table A3b and not reported					
Ideological imprint (H1)	-.266**	-.014*	-.294**	-.028**	-.377**	-.040**	-2.321**	-.228**	-2.633**	-.269**	-3.284**	-.319**
	(.063)	(.006)	(.077)	(.010)	(.080)	(.005)	(.559)	(.048)	(.520)	(.057)	(.621)	(.059)
Political involvement	.088	.001	.113	.009**	.032	.012**	.259	.095	.002	.000	.494**	.120**
	(.288)	(.006)	(.144)	(.003)	(.680)	(.004)	(.393)	(.071)	(.083)	(.003)	(.058)	(.012)
Government appropriation	.601	.003	8.871	.017	2.503	.038**	1.469	.118	.026	.014	5.203*	.047
	(1.176)	(.029)	(5.840)	(.137)	(2.007)	(.004)	(3.651)	(.329)	(.255)	(.011)	(2.409)	(.049)
Social network	.378	.010**	.062*	.034*	.310	.002	4.243	.095	.002	.000	.008	.001
	(.273)	(.003)	(.026)	(.014)	(.838)	(.004)	(5.198)	(.067)	(.031)	(.001)	(.012)	(.003)
Post 2001	.008	.000	.015**	.003**	.008	.000	.011	.003*	.006**	.000 ⁺	.021**	.004**
	(.006)	(.000)	(.005)	(.001)	(.013)	(.000)	(.009)	(.001)	(.002)	(.000)	(.002)	(.000)
Regional FDI intensity	1.346	.009	.083**	.228**	5.580	.029*	.258**	.231 ⁺	.549**	.012*	.051**	.406**
	(.879)	(.009)	(.031)	(.043)	(8.129)	(.012)	(.082)	(.124)	(.195)	(.006)	(.019)	(.114)
Ideological imprint × Political involvement (H2a)	.189**	.010*	.126**	.095**	.477**	.017**	1.568*	.172*	1.973**	.140**	1.580 ⁺	.066**
	(.061)	(.005)	(.045)	(.031)	(.141)	(.005)	(.682)	(.081)	(.740)	(.037)	(.949)	(.016)
Ideological imprint × Government appropriation (H2b)	.889*	.093*	.530*	.523**	2.481**	.400**	3.766**	.319**	5.096**	.813**	4.647 ⁺	.428**
	(.365)	(.043)	(.245)	(.203)	(.878)	(.058)	(1.057)	(.050)	(1.587)	(.228)	(2.683)	(.096)
Ideological imprint × Social network (H3)	.071*	.005**	.059*	.049**	.240**	.006**	.510**	.095**	.931**	.061**	.651**	.046*
	(.025)	(.001)	(.027)	(.016)	(.063)	(.002)	(.087)	(.011)	(.344)	(.009)	(.122)	(.019)
Ideological imprint × Post-1999 (H4a)	.015*	.003**	.011*	.048*	.125*	.003*	.139*	.055**	.339*	.017**	.185**	.012**
	(.007)	(.001)	(.005)	(.022)	(.053)	(.001)	(.060)	(.014)	(.180)	(.006)	(.044)	(.002)
Ideological imprint × Regional FDI intensity (H4b)	.041**	.006**	.011 ⁺	.026**	.181**	.005**	.357*	.069**	.709**	.014 ⁺	.314**	.001**
	(.007)	(.001)	(.006)	(.006)	(.056)	(.001)	(.161)	(.018)	(.268)	(.008)	(.058)	(.000)
Number of observations	4,483	679	1,686	255	14,395	2,181	3,078	540	1,158	203	9,883	1,733
Goodness of fit (pseudo R ²)	.377		.395		.386		.453		.432		.454	
Goodness of fit (between R ²)		.311		.365		.297		.372		.401		.384

+ $p < .10$; * $p < .05$; ** $p < .01$; intercept not reported to save space.

* Numbers of observation vary according to different subsamples.

† B: binary measure of dependent variable, C: continuous measure of dependent variable.

‡ Other industries include services, advertisement, consulting, art, retail, etc.

REFERENCES

Abadie, A.

2005 "Semiparametric difference-in-differences estimators." *Review of Economic Studies*, 72: 1–19.

Andersen, P. K.

1982 "Testing goodness of fit of Cox's regression and life model." *Biometrics*, 38: 67–77.

Bian, Y., X. Shu, and J. R. Logan

2001 "Communist Party membership and regime dynamics in China." *Social Forces*, 79: 805–841.

Blevins, D. P., E. W. Tsang, and S. M. Spain

2015 "Count-based research in management: Suggestions for improvement." *Organizational Research Methods*, 18: 47–69.

Hausman, J., B. H. Hall, and Z. Griliches

1984 "Econometric models for count data with an application to the patents–R&D relationship." *Econometrica*, 52: 909–938.

Hausman, J. A.

1978 "Specification tests in econometrics." *Econometrica*, 46: 1251–1271.

Jia, N.

2014 "Are collective political actions and private political actions substitutes or complements? Empirical evidence from China's private sector." *Strategic Management Journal*, 35: 292–315.

Jost, J. T., C. M. Federico, and J. L. Napier

2009 "Political ideology: Its structure, functions, and elective affinities." *Annual Review of Psychology*, 60: 307–337.

Kung, J. K.-s., and S. Chen

2011 "The tragedy of the nomenklatura: Career incentives and political radicalism during China's Great Leap famine." *American Political Science Review*, 105: 27–45.

Lin, D. Y., and L.-J. Wei

1989 "The robust inference for the Cox proportional hazards model." *Journal of the American Statistical Association*, 84: 1074–1078.

Liu, X., T. Buck, and C. Shu

2005 "Chinese economic development, the next stage: Outward FDI?" *International Business Review*, 14: 97–115.

Lu, X.

2004 *Rhetoric of the Chinese Cultural Revolution: The Impact on Chinese Thought, Culture, and Communication*. Columbia, SC: University of South Carolina Press.

Luo, Y., Q. Xue, and B. Han

2010 "How emerging market governments promote outward FDI: Experience from China." *Journal of World Business*, 45: 68–79.

Marquis, C., and A. Tilcsik

2013 "Imprinting: Toward a multilevel theory." *Academy of Management Annals*, 7: 195–245.

Marquis, C., and A. Tilcsik

2016 "Institutional equivalence: How industry and community peers influence corporate philanthropy." *Organization Science*, 27: 1325–1341.

Schoenfeld, D.

1982 "Partial residuals for the proportional hazards regression model." *Biometrika*, 69: 239–241.

Villalonga, B., and A. M. McGahan

2005 "The choice among acquisitions, alliances, and divestitures." *Strategic Management Journal*, 26: 1183–1208.

Wu, X., and D. J. Treiman

2007 "Inequality and equality under Chinese socialism: The hukou system and intergenerational occupational mobility." *American Journal of Sociology*, 113: 415–445.

Zellner, A.

1962 "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias." *Journal of the American Statistical Association*, 57: 348–368.

Zhang, J., C. Marquis, and K. Qiao

2016 "Do political connections buffer firms from or bind firms to the government? A study of corporate charitable donations of Chinese firms." *Organization Science*, 27: 1307–1324.