

Supplementary Materials

Does the County-Level Business Environment Improve the Quality of Agriculture-Related Market Entry? Evidence from China

Supplementary Note

These supplementary materials provide additional information on text-based identification rules, variable construction, data processing, and supporting robustness checks. They are intended to improve transparency and facilitate editorial and reviewer verification. The tables are supplementary to, and do not replace, the definitions, empirical models, and main results reported in the manuscript.

Table S1. Representative keyword dictionary for identifying agriculture-related value chain activities.

Identification category	Screening keywords	Description	Rationale
Agricultural product processing	agricultural product processing; grain processing; fruit and vegetable processing; tea processing; meat-product processing; aquatic-product processing	Processing of agricultural, grain, fruit, vegetable, tea, meat, and aquatic products.	Captures value-added activities beyond primary production.
Cold chain storage and logistics	cold chain; refrigerated storage; cold warehouse; preservation; warehousing; logistics; distribution; transportation	Cold chain, warehousing, logistics, distribution, and transportation services.	Identifies circulation services that reduce spoilage and spatial mismatch.
Agricultural inputs and machinery services	seeds; seedlings; fertilizer; pesticides; feed; agricultural machinery; machinery services; machinery leasing	Agricultural inputs, machinery services, and machinery leasing.	Reflects upstream input supply and mechanized services.
Agricultural technology services	agricultural technology; technology extension; planting technology; breeding technology; agricultural consulting	Technology extension, planting and breeding technology, and agricultural consulting.	Captures knowledge-intensive upgrading services.
E-commerce and circulation	e-commerce; online sales; agricultural product sales; wholesale; retail; supply chain; circulation services	E-commerce, online sales, wholesale, retail, supply chain, and circulation services.	Identifies activities linking county producers with wider markets.
Quality testing and certification	quality testing; inspection; certification; traceability; food safety; standardization	Quality testing, inspection, certification, traceability, food safety, and standardization.	Represents quality control and compliance services.
Branding and marketing services	brand; packaging; marketing; exhibition; promotion; product design; regional public brand	Branding, packaging, marketing, exhibitions, promotion, and product design.	Captures downstream market development and differentiation.
Other value chain services	agricultural socialized services; trusteeship services; leisure agriculture; rural tourism; agricultural information services; supply and marketing services	Agricultural socialized services, trusteeship, leisure agriculture, rural tourism, information services, and supply marketing services.	Identifies service-oriented entrants supporting sectoral integration.

Note: Table S1 reports the representative keyword dictionary used for business scope screening. Business scope strings were standardized before matching. Firms classified in agriculture, forestry, animal husbandry, and fishery were retained under the narrow-scope definition; the dictionary identifies broader agricultural value chain activities.

Table S2. Variable construction, data level, and preprocessing rules.

Variable or item	Construction	Data source/level	Preprocessing and use in estimation
High-quality entry (HQE)	Composite county-year measure based on capital quality, organizational quality, and value chain extension among newly registered agriculture-related entities.	Firm registration records aggregated to the county-year level.	Bounded between 0 and 1. Component-level checks are reported in Table S4.
Capital quality component	Measures registered capital capacity among newly registered agriculture-related entities.	Firm registration records, county-year aggregation.	Used as one component of HQE and as a separate dependent variable in the component check.
Organizational quality component	Measures whether entrants adopt more formalized organizational forms.	Firm registration records, county-year aggregation.	Captures governance and organizational upgrading.
Value chain extension component	Measures involvement in processing, circulation, technology, services, or other non-primary value chain activities.	Firm industry classification and business scope text, county-year aggregation.	Keyword matching follows the representative dictionary in Table S1.
County-level business environment index	Entropy-weighted composite index covering government services, market environment, finance, human capital, public services, and infrastructure.	County-level socioeconomic statistics and county-year panel indicators.	Positive indicators are standardized before aggregation. Full-sample fixed entropy weights improve county-year comparability.
Government services	General public budget revenue/GDP; per capita general public budget expenditure.	County-year statistical indicators.	Included as one dimension of the business environment index.
Market environment	Per capita GDP; tertiary industry value added/GDP; industrial firms above designated size per 10,000 persons; per capita retail sales of consumer goods.	County-year statistical indicators.	Included in the index; a robustness check excludes this market dimension.
Financial environment	Financial institution loan balance/GDP; financial institution loan balance/resident savings.	County-year statistical indicators.	Used in the index and pathway analysis.
Human capital environment	Students in primary and secondary schools per 1,000 persons.	County-year statistical indicators.	Used in the index and pathway analysis.
Public services	Hospital and health center beds per 10,000 persons; social welfare institutions per 10,000 persons.	County-year statistical indicators.	Included as one dimension of the business environment index.
Infrastructure	Fixed-line telephone users/total households.	County-year statistical indicators.	Included as one dimension of the business environment index.
Baseline sample rule	County-year observations with fewer than 10 newly registered agriculture-related entrants are excluded.	County-year panel constructed from firm registration records and county statistics.	This mitigates volatility from very small denominators. Threshold checks of at least 5 and at least 20 entrants are reported in Table S3.

Note: This table is designed for replicability and reviewer verification. It summarizes the construction and processing logic underlying the variables used in the manuscript. More detailed econometric specifications are reported in the main text.

Table S3. Supporting robustness and sample threshold sensitivity checks.

Check	Estimated coefficient	Std. error	p-value	Observations	R-squared
Equal-weighted business environment index	0.3233***	0.0543	0.0000	32,347	0.0196
Full-indicator entropy index	0.5287***	0.0825	0.0000	32,347	0.0198
Average quality score	0.6272***	0.1052	0.0000	32,347	0.0134
Log number of high-quality entrants	1.3311***	0.1710	0.0000	32,347	0.0218
Winsorized dependent variable	0.4146***	0.0631	0.0000	32,347	0.0209
Excluding municipal districts	0.4913***	0.0946	0.0000	21,516	0.0186
Entrants no fewer than 20	0.3906***	0.0637	0.0000	31,919	0.0213
Current explanatory variable	0.3270***	0.0627	0.0000	34,999	0.0191
Two-period lagged explanatory variable	0.4259***	0.0642	0.0000	29,689	0.0200
Province-year fixed effects	0.1152*	0.0641	0.0723	32,347	0.0109
Prefecture-year fixed effects	0.0602	0.0624	0.3351	32,082	0.0071
Narrow-scope agricultural firms	0.3588***	0.0669	0.0000	31,921	0.2103
Fractional logit marginal effect	0.4520***	0.0663	0.0000	32,351	
Fractional probit marginal effect	0.4544***	0.0661	0.0000	32,351	
Institutional service index excluding market dimension	0.2457***	0.0467	0.0000	32,351	0.0189
Entrants no fewer than 5	0.4310***	0.0627	0.0000	32,608	0.0204

Note: The table reports robustness checks referenced in the manuscript. Standard errors are clustered at the county level unless otherwise stated in the main text. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. R-squared is not reported for fractional response models because they are nonlinear fractional response specifications.

Table S4. Component-level checks of high-quality entry.

Dependent variable	Estimated coefficient	Std. error	p-value	Observations	R-squared
Capital quality component	0.2445***	0.0524	0.0000	32,351	0.0035
Organizational quality component	0.3140***	0.0616	0.0000	32,351	0.0199
Value chain extension component	0.0625*	0.0380	0.0999	32,351	0.0046

Note: Each row uses one component of high-quality entry as the dependent variable. The specification follows the baseline fixed-effects framework in the manuscript. Standard errors are clustered at the county level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.