Proceedings of ICEMGD 2026 Symposium: Green Economy and Industrial Ethics: Sectoral Innovation of TCM Derivatives
DOI: 10.54254/2754-1169/2026.HZ36743

Conditional Effects of Digital Financial Inclusion Policies on Rural Household Entrepreneurship Evidence from Asset Constraints, Moderated Difference-in-Differences, and Causal Forests

Xingyou Li

Department of Dual Degree in Statistics and Finance, Central University of Finance and Economics, Beijing, China
3881348672@qq.com

Abstract. This study examines which rural households translate digital financial inclusion policies into entrepreneurial action. A balanced 2010–2022 China Family Panel Studies panel is linked to provincial digital finance, digital infrastructure, policy-pilot, and macroeconomic data, yielding a 2012–2022 rural sample. The empirical strategy combines a pre-policy asset-constraint moderated difference-in-differences design, an event study, and an honest causal forest. Average effects of the aggregate digital financial inclusion index and binary policy exposure are unstable. Policy intensity, measured by the number of active related pilots in each province-year, significantly increases entrepreneurship among households with low baseline assets: each additional pilot raises their entrepreneurship probability by about 1.85 percentage points, while the effect for higher-asset households is insignificant. The result survives a stricter entrepreneurship definition and the exclusion of centrally administered municipalities. Post-policy effects for low-asset households reach 5.38–5.74 percentage points, and the complete effect is 7.63 percentage points for low-asset households with baseline internet access. Causal-forest profiles identify younger, healthier, digitally connected households with greater labor supply as high responders. The findings support targeted policy expansion centered on asset constraints and minimum digital access.

Keywords: digital financial inclusion policy, rural household entrepreneurship, asset constraints, moderated difference-in-differences, causal forest

1. Introduction and literature review

1.1. Background and research questions

Rural household entrepreneurship supports employment, operating income, and county-level economic vitality, yet rural households often face limited collateral, incomplete credit records, sparse financial outlets, and weak market information. Digital financial inclusion can reduce these

DOI: 10.54254/2754-1169/2026.HZ36743

barriers through mobile payments, online credit, digital accounts, and data-based risk assessment, making it a potentially important instrument of rural revitalization.

Digital financial supply, however, does not automatically generate entrepreneurship. Asset shortages, limited risk-bearing capacity, weak digital skills, and poor links between finance and productive opportunities may prevent households from using new services. This study therefore asks whether related policies benefit resource-constrained rural households and which household conditions allow policy support to become entrepreneurial action.

The question is policy-relevant because effective targeting requires simultaneous attention to financing constraints and digital accessibility rather than account coverage alone.

1.2. Literature review

The literature falls into four related strands.

First, Guo et al. [1] developed the multilevel Peking University Digital Financial Inclusion Index, measuring coverage breadth, usage depth, and digitization and documenting substantial spatial inequality in digital finance.

Second, studies of inclusiveness report mixed effects. Zhang et al. [2] linked the index to China Family Panel Studies (CFPS) data and found gains in rural income and entrepreneurial opportunities; Song [3] reported a narrower urban-rural income gap, whereas He and Wang [4] showed that the digital divide may widen it in some regions. Thus, benefits depend on households' and regions' capacity to absorb digital services.

Third, mechanism studies emphasize payment convenience and liquidity relief [5], heterogeneous effects across entrepreneurial types [6], and limits in reducing entry costs [7]. These findings imply that financial access, digital access, resource constraints, and entrepreneurship form a complex chain that is unlikely to be captured by a single mediator.

Fourth, recent work combines quasi-natural experiments with machine learning. Li et al. [8] used reform pilots and double machine learning, Xie et al. [9] examined heterogeneous policy effects, and Wager and Athey [10] developed causal forests for treatment-effect heterogeneity. This literature motivates a shift from unstable average effects to the conditions under which policy works.

1.3. Research positioning and marginal contributions

This study makes three contributions. It identifies conditional rather than universal policy effects by treating pre-policy asset constraints as the main moderator. It avoids overcontrol by excluding current income, assets, debt, and internet use from the total-effect specification when they may be policy outcomes. Finally, moderated DID provides the main inference, while an honest causal forest profiles high-CATE households and an interaction DID supplies an interpretable cross-check.

2. Theoretical analysis and research hypotheses

2.1. Policy compensation under asset constraints

Entrepreneurship requires start-up and working capital and the capacity to bear risk. Higher-asset households may already rely on self-financing, formal credit, or informal loans, so additional policy support has limited marginal value. For low-asset households, financing constraints are more binding. By expanding microcredit, credit information, payments, and service coordination, inclusive-finance pilots should therefore generate a stronger compensatory effect for this group.

H1: The entrepreneurial effect of related policy intensity is stronger for rural households with low baseline assets than for higher-asset households, conditional on household fixed effects and year shocks.

2.2. Minimum condition of digital access

Digital financial services can affect entrepreneurship only if households can access digital tools. Internet access is not advanced digital capability, but it provides a basic channel for policy information, online payments, digital credit, and online markets. Without this gateway, formal policy availability may not become effective access.

H2: Among low-asset households, baseline internet access strengthens the positive effect of policy intensity on entrepreneurship.

2.3. From average effects to policy-response profiles

Responses may still differ by age, health, labor supply, education, and local digital conditions. A causal forest can identify multidimensional CATE profiles beyond prespecified subgroup regressions. Here it is used to characterize heterogeneity and generate testable profiles, not to replace moderated DID.

3. Research design

3.1. Data sources and sample

The analysis combines four data sources: the 2010–2022 balanced CFPS household panel; the 2011–2023 provincial Peking University Digital Financial Inclusion Index (DPFI); inclusive-finance, rural-revitalization finance, and related reform-pilot records aggregated by province-year; and provincial macroeconomic and digital-infrastructure indicators.

Because the DPFI begins in 2011, the main sample retains rural observations from 2012, 2014, 2016, 2018, 2020, and 2022. It contains 11,456 household-year observations for 2,197 households in 29 provincial-level regions. Entrepreneurship equals one when a household reports industrial, commercial, or entrepreneurial activity. Since the earliest pilot began in 2015, baseline assets, internet access, age, education, and health are averaged over 2012–2014 to avoid post-policy classification.

3.2. Core variables and treatment principles

The dependent variable is household entrepreneurship. BroadPost indicates whether at least one related pilot is active in a province-year. Intensity counts distinct active pilots. LowAsset equals one when pre-policy net assets are below the sample median, and Internet records baseline household internet access.

Intensity is an active-pilot count covering inclusive-finance reform zones, rural-revitalization inclusive-finance zones, comprehensive rural finance reforms, and micro- and small-enterprise financial-service reforms. Each verified pilot is active from its approval year and, absent an announced end date, through the sample period. Concurrent pilots are counted separately, as formalized in equation (1).

DOI: 10.54254/2754-1169/2026.HZ36743

$$Intensity_{pt} = \sum_{k=1}^{K} 1(p \in S_k, t \geq T_k) \quad (1)$$

In equation (1), S_k is the set of provinces covered by pilot k and T_k is its start year. Intensity therefore ranges from 0 to 3, the observed maximum rather than an imposed cap. Zhejiang illustrates the coding: the Taizhou pilot yields Intensity = 1 in 2015–2018; adding the Ningbo inclusive-finance pilot raises it to 2 in 2019–2021; adding the Lishui rural-revitalization pilot raises it to 3 in 2022. Equal weighting makes the variable a measure of overlapping institutional supply, not fiscal scale, implementation quality, or the strength of individual instruments.

The total-effect model controls for demographics unlikely to be policy outcomes—head age, household size, sex, marital status, medical insurance, and health—and lagged provincial digital infrastructure. Current income, assets, debt, internet use, and the DPFI may lie on the transmission path and are excluded to avoid overcontrol. The aggregate DPFI appears only in the average-effect diagnostic.

Table 1. Descriptive statistics of the main variables

Variable	Observations	Mean	Std. dev.	Median	Minimum	Maximum
Household entrepreneurship	11,456	0.069	0.254	0.000	0.000	1.000
Policy-pilot intensity	11,456	0.157	0.407	0.000	0.000	3.000
Low baseline assets	11,456	0.493	0.500	0.000	0.000	1.000
Baseline internet access	11,456	0.308	0.462	0.000	0.000	1.000
Young household head at baseline	11,456	0.478	0.500	0.000	0.000	1.000
Aggregate DPFI (standardized)	11,456	−0.110	0.882	−0.134	−1.606	1.855
Degree of digitization (standardized)	11,456	−0.033	0.859	0.255	−1.818	1.242
Baseline household net assets (standardized)	11,298	−0.134	0.718	0.014	−5.869	0.927
Baseline per capita net income (standardized)	11,298	−0.591	0.813	−0.537	−3.053	2.039
Baseline years of education (standardized)	11,298	−0.401	0.848	−0.329	−1.612	1.809
Baseline digital infrastructure (standardized)	11,298	−0.390	0.421	−0.561	−1.085	1.348

Note: Rural observations cover 2012–2022. Baseline variables use 2012–2014 information. Policy-pilot intensity counts simultaneously active related pilots in each province-year and ranges from 0 to 3.

DOI: 10.54254/2754-1169/2026.HZ36743

3.3. Average-effect diagnostics and moderated DID model

The analysis first diagnoses average effects of the aggregate DPFI and policy variables, then estimates the moderated DID model in equation (2).

$$Entrepreneur_{ipt} = \alpha + \beta \cdot Intensity_{pt} + \theta \cdot (Intensity_{pt} \times LowAsset_i^0) + \gamma\prime X_{it} + \quad (2)$$

$$\mu_i + \lambda_t + \varepsilon_{ipt}$$

Y is household entrepreneurship, D policy intensity, L^0 the low-asset indicator, and i, p, and t index households, provinces, and years. Household and year fixed effects are μ_i and λ_t; X contains predetermined demographics and lagged digital infrastructure. The coefficient θ captures the additional effect for low-asset households. Standard errors are clustered by province. Dynamic specifications interact LowAsset with distant pre-policy, 0–1-year post-policy, and 2+-year post-policy indicators; each low-asset conditional effect is the sum of the main and interaction terms.

3.4. Causal forest and fact-based cross-check

Within the low-asset subsample, CausalForestDML estimates the conditional marginal effect of one additional unit of policy intensity, as shown in equation (3).

$$\tau(x) = \partial E[Entrepreneur_{ipt} \vee Intensity_{pt}, X_i^0 = x, W_{ipt}]/\partial Intensity_{pt} \quad (3)$$

Baseline features include pre-policy income, assets, education, internet access, health, age, household size, and digital infrastructure; nuisance controls also include province and year indicators. The model uses 800 trees, honest splitting, and five-fold cross-fitting. Because treatment varies at the province-year level and the number of provinces is limited, CATEs are used only for profiling. The internet-access profile identified by the forest is then tested in an interaction DID.

3.5. Identification boundaries and robustness strategy

The DID design uses provincial exposure although some pilots operate at the city or county level, which may dilute treatment. Estimates therefore represent the conditional effect of greater province-level policy supply, not a pilot-city local effect. The limited number of provincial clusters also requires cautious inference. Robustness checks use a stricter entrepreneurship outcome and exclude Beijing, Tianjin, Shanghai, and Chongqing.

4. Empirical results

4.1. Average-effect diagnostics: no stable universal promotion effect

Table 2 shows no stable universal effect. The aggregate DPFI estimate is −0.0101 and insignificant; broad policy exposure and policy intensity are positive but not robustly significant with predetermined controls. Average coefficients can therefore conceal heterogeneous responses across household resource positions.

DOI: 10.54254/2754-1169/2026.HZ36743

Table 2. Average-effect diagnostics for digital financial inclusion and policy variables

Model	Core variable	Coefficient (standard error)	p value	Observations	Fixed effects
(1) Aggregate DPFI	dpfi_total_z	−0.0101(0.0321)	0.753	11,456	Household, year
(2) Broad policy exposure	pfi_broad_post	0.0108(0.0074)	0.146	11,456	Household, year
(3) Policy-pilot intensity	policy_intensity	0.0083(0.0051)	0.102	11,456	Household, year

Note: Models (2) and (3) control for predetermined demographics and lagged digital infrastructure but exclude current income, assets, debt, internet use, and the DPFI as potential policy outcomes.

4.2. Main result: asset constraints significantly moderate the entrepreneurial effect of policy intensity

Table 3 reports the main result. Policy intensity has essentially no effect for higher-asset households, whereas its interaction with low baseline assets is positive in all specifications. In column (3), the interaction is 0.0179 ($p = 0.002$) and the complete effect for low-asset households is 0.0185 ($p = 0.006$). Thus, each additional active pilot raises their entrepreneurship probability by about 1.85 percentage points. H1 is supported.

Table 3. Moderating role of asset constraints in the entrepreneurial effect of policy intensity

Variable	(1) Household and year FE	(2) + demographic controls	(3) + lagged digital infrastructure
Policy intensity (higher-asset households)	−0.0001(0.0056)	0.0008(0.0054)	0.0006(0.0054)
Policy intensity × low baseline assets	0.0180***(0.0063)	0.0181***(0.0059)	0.0179***(0.0059)
Conditional effect for low-asset households	0.0179***(0.0069)	0.0189***(0.0067)	0.0185***(0.0067)
Observations	11,456	11,456	11,456
R^2	0.000	0.004	0.004

Note: Low baseline assets are defined by the median of average household net assets in 2012–2014. Province-clustered standard errors are in parentheses; *, **, and *** denote significance at the 10%, 5%, and 1% levels.

DOI: 10.54254/2754-1169/2026.HZ36743

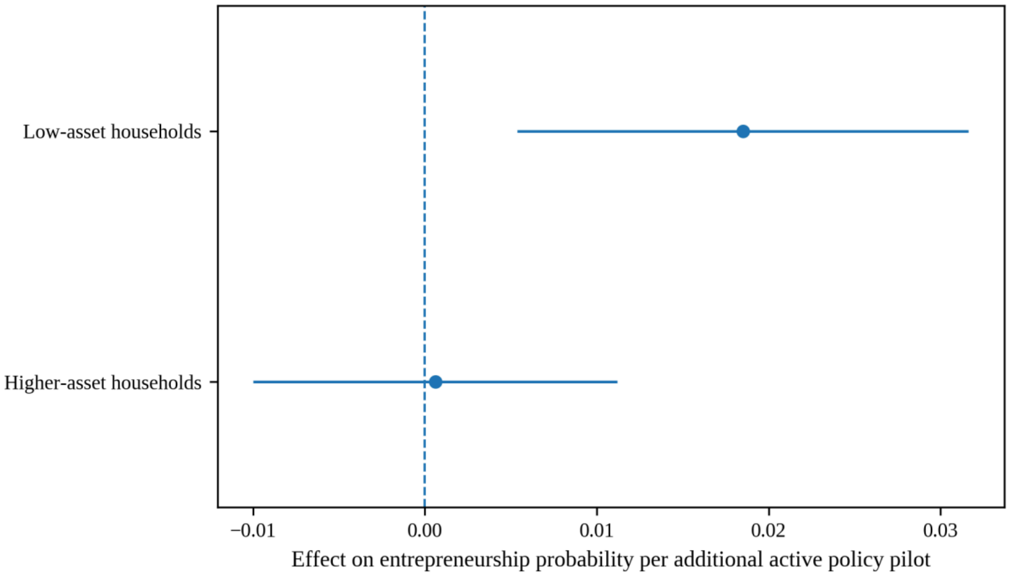

Figure 1. Conditional effect of policy intensity on entrepreneurship by household asset status

Two robustness checks retain the column (3) specification. The first defines strict entrepreneurship as reported industrial, commercial, or entrepreneurial activity with positive productive fixed assets. The second excludes Beijing, Tianjin, Shanghai, and Chongqing to reduce the influence of their distinctive urbanization, financial concentration, and policy transmission.

Table 4. Robustness checks using an alternative outcome and excluding municipalities

Variable	(1) Baseline	(2) Strict entrepreneurship	(3) Excluding municipalities
Policy intensity (higher-asset households)	0.0006(0.0054)	−0.0003(0.0054)	0.0003(0.0055)
Policy intensity × low baseline assets	0.0179*** (0.0059)	0.0180***(0.0061)	0.0184***(0.0058)
Conditional effect for low-asset households	0.0185*** (0.0067)	0.0177***(0.0067)	0.0187***(0.0068)
Observations	11,456	11,456	11,172
Provincial clusters	29	29	25

Note: Column (1) reproduces Table 3, column (3). Column (2) uses strict entrepreneurship; column (3) excludes the four municipalities. All models include household and year fixed effects, demographics, and lagged digital infrastructure. Province-clustered standard errors are in parentheses.

Table 4 confirms robustness. Under the strict outcome, the interaction and complete low-asset effect are 0.0180 and 0.0177; after excluding municipalities, they are 0.0184 and 0.0187. All remain significant at 1% and close to the baseline estimates.

DOI: 10.54254/2754-1169/2026.HZ36743

4.3. Dynamic test: post-policy effects for low-asset households

Table 5 and Figure 1 report event-time effects for low-asset households. The distant pre-policy coefficient is 0.0399 and not significant at 5%. Effects rise to 0.0538 ($p = 0.040$) in years 0–1 and 0.0574 ($p = 0.027$) after two years. Because the pre-policy estimate approaches the 10% threshold and exposure is matched provincially, the evidence supports but does not conclusively establish parallel trends.

Table 5. Dynamic policy effects for low-asset households

Relative period	Conditional effect	Standard error	p value	Interpretation
Distant pre-policy period	0.0399	(0.0248)	0.107	Parallel-trends diagnostic
0–1 years after policy	0.0538**	(0.0261)	0.040	Short-run policy response
2+ years after policy	0.0574**	(0.0259)	0.027	Persistent policy response

Note: Conditional effects sum the event-time main effect and its interaction with low assets; the near pre-policy period is the reference. Controls match the main specification.

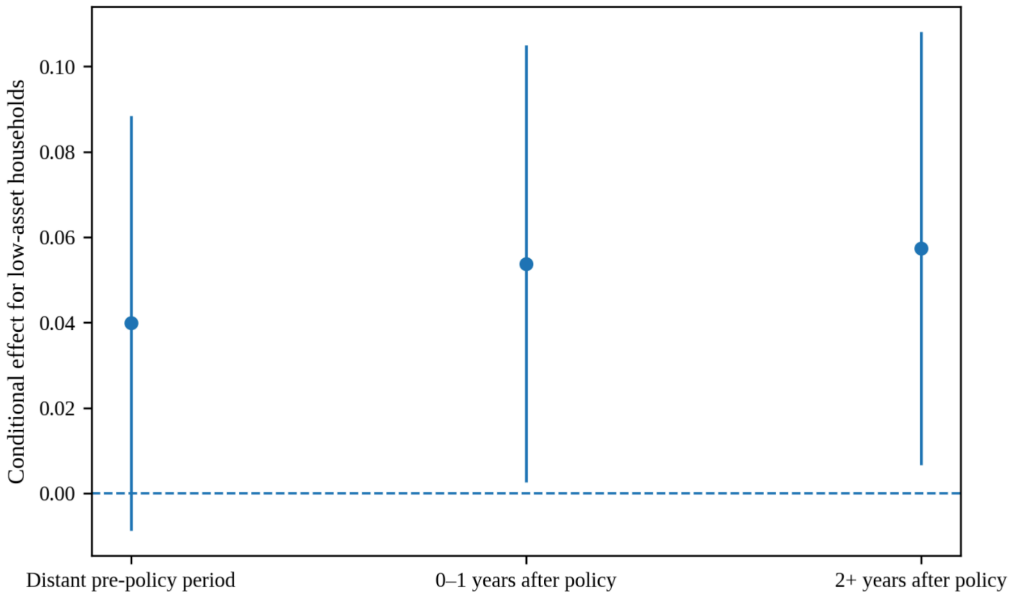

Figure 2. Dynamic policy effects for low-asset households with 95% confidence intervals

4.4. Fact-based cross-check: minimum digital access as the key condition for policy translation

The causal forest suggests that high-CATE low-asset households have greater baseline internet access. A full-sample triple interaction tests this profile. In Table 6, policy intensity × low assets × internet access equals 0.0712 ($p = 0.068$), while the complete effect for low-asset, internet-connected households is 0.0763 ($p = 0.040$). Other asset–internet groups show no significant effect. H2 therefore receives preliminary support.

DOI: 10.54254/2754-1169/2026.HZ36743

Table 6. Triple-moderation test for low assets and internet access

Variable	Coefficient (standard error)	p value
Policy intensity (higher assets, no internet: reference)	0.0001(0.0101)	0.992
× low assets	0.0037(0.0094)	0.696
× internet access	0.0013(0.0230)	0.953
× low assets × internet access	0.0712*(0.0390)	0.068
Complete effect: low assets with internet access	0.0763**(0.0371)	0.040

Note: Baseline internet access is averaged over 2012–2014. The complete effect is the linear combination of policy intensity and all relevant interactions.

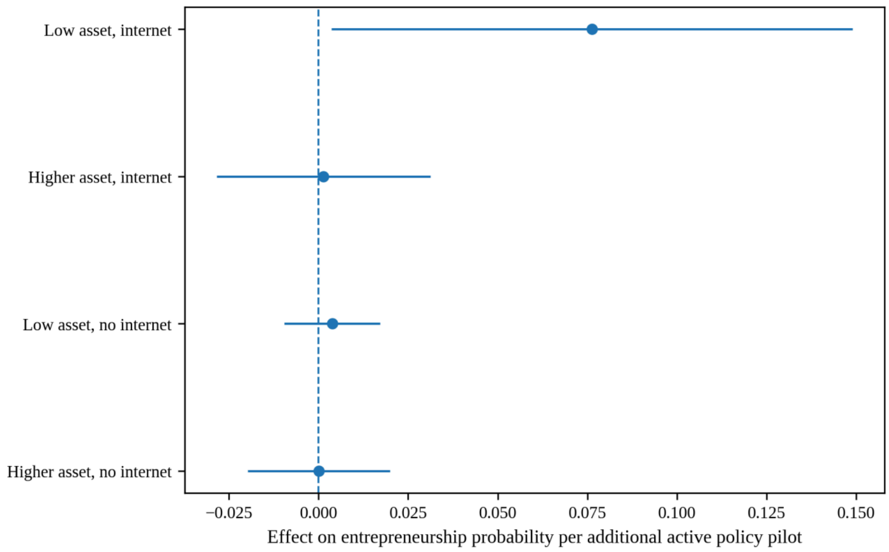

Figure 3. Conditional effects of policy intensity across asset–internet groups

4.5. Causal-forest CATE profiles: which households are highly responsive

Within the low-asset subsample, the honest causal forest estimates a mean CATE of 0.0147 per additional pilot. The highest-CATE tercile averages 0.0387, compared with −0.0022 in the lowest tercile, indicating substantial heterogeneity. Because these are predictions, interpretation relies on profiles and the interaction-DID cross-check rather than conventional significance stars.

Table 7. CATE distribution among low-asset households

CATE group	Observations	Mean CATE	Median CATE	Share with positive effect
All low-asset observations	5,649	0.0147	0.0070	0.770
Lowest-CATE tercile	1,867	−0.0022	−0.0013	0.305
Highest-CATE tercile	1,867	0.0387	0.0332	1.000

Note: CATE is the predicted conditional marginal effect of one additional unit of policy intensity and is used for heterogeneity profiling, not as a substitute for moderated DID.

DOI: 10.54254/2754-1169/2026.HZ36743

Table 8. Baseline characteristics of high- and low-CATE low-asset households

Characteristic	Low-CATE group	High-CATE group	Standardized difference (high − low)
Baseline internet access	0.052	0.341	0.950
Household-head age (standardized)	0.597	−0.586	−1.366
Health status (standardized)	−0.652	−0.007	0.769
Household size	2.983	4.466	0.859
Years of education (standardized)	−0.581	−0.456	0.145
Baseline digital infrastructure (standardized)	−0.243	−0.641	−1.049

Note: High- and low-CATE groups are the top and bottom terciles within the low-asset sample. Internet access is binary; other entries are interpreted using the original variable direction.

High-CATE households have a baseline internet-access rate of 34.1%, versus 5.2% in the low-CATE group, and have younger, healthier heads and larger households. Their regions do not have better baseline digital infrastructure. Policy benefits may therefore be especially compensatory where households possess basic internet access but regional digital development remains incomplete. Figure 2 summarizes these standardized differences.

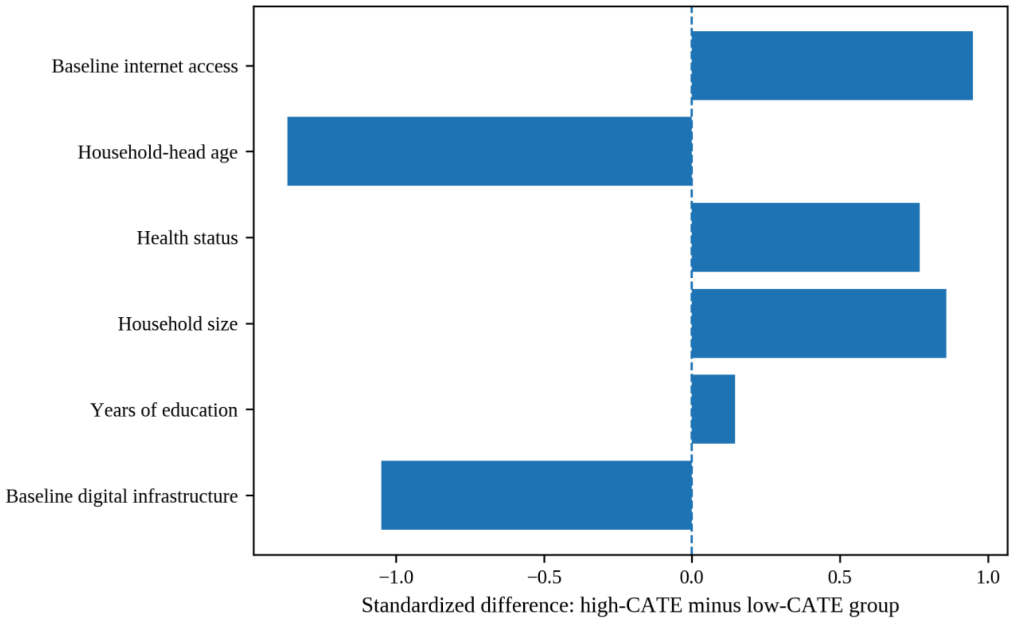

Figure 4. Standardized baseline differences between high- and low-CATE low-asset households

4.6. Summary of results

The evidence is coherent but conditional: average effects are unstable; policy intensity compensates low-asset households and survives alternative outcomes and samples; internet access is a key threshold; and the causal forest identifies digitally connected, younger, healthier low-asset

DOI: 10.54254/2754-1169/2026.HZ36743

households as high responders. Policy operates at the intersection of resource constraints and minimum digital access rather than through universal coverage expansion.

5. Conclusions and recommendations

5.1. Conclusions

This study reassesses the entrepreneurial effects of digital financial inclusion policies using a rural CFPS panel linked to provincial policy and contextual data. The aggregate DPFI and average policy measures do not establish a universal effect. Policy intensity, however, significantly raises entrepreneurship among low-baseline-asset households, remains robust to a stricter outcome and the exclusion of municipalities, and produces stronger post-policy responses. The largest effect occurs among low-asset households with internet access. Causal-forest profiles further identify younger, healthier, digitally connected households with greater labor supply as high responders.

5.2. Policy recommendations

First, targeting should prioritize asset-constrained households. Financial institutions and local governments can combine compliant information on agricultural operations, public payments, input purchases, and rural e-commerce to offer small, staged, purpose-matched financing to households with entrepreneurial potential but little collateral.

Second, minimum digital access should be treated as a policy prerequisite. Smart devices, stable networks, basic digital training, and secure online payments should be coordinated with credit, guarantees, and entrepreneurship support.

Third, policy supply should be assessed by the density of complementary instruments rather than pilot designation alone. Inclusive-finance reform, rural e-commerce, training, risk compensation, and industrial-chain matching should be integrated to reduce policy fragmentation.

Fourth, high-response profiles can inform service sequencing but should not replace human assessment. Low-asset households with internet access and adequate labor may be reached first, while households lacking connectivity or facing health and labor constraints require more basic public services to prevent wider entry gaps.

5.3. Limitations and future research

Four limitations remain. Provincial matching may dilute city- or county-level pilots; the limited number of provincial clusters warrants cautious inference; equal-weighted pilot counts capture overlapping policy supply but not spending, implementation quality, or instrument differences; and causal-forest CATEs provide profiles rather than independent causal estimates. Future research should use finer geographic identifiers, longer pre-policy periods, direct credit and payment measures, and outcomes distinguishing entrepreneurial type, persistence, scale, and income quality.

References

[1] Guo, F., Wang, J., Wang, F., et al. (2020). Measuring China's digital financial inclusion: Index compilation and spatial characteristics. *China Economic Quarterly, 19*(4), 1401–1418.

[2] Zhang, X., Wan, G., Zhang, J., et al. (2019). Digital economy, financial inclusion, and inclusive growth. *Economic Research Journal, 54*(8), 71–86.

[3] Song, X. (2017). An empirical test of whether digital financial inclusion narrows the urban-rural income gap. *Finance & Economics*, (6), 14–25.

DOI: 10.54254/2754-1169/2026.HZ36743

[4] He, Z., & Wang, Q. (2026). Effects of digital financial inclusion on the urban-rural income gap: Evidence from a two-way fixed-effects spatial Durbin model. *China Journal of Commerce, 35*(2), 105–109.
[5] Zhang, L., & Wen, T. (2022). How does digital financial inclusion affect rural industrial integration? *Chinese Rural Economy*, (7), 59–80.
[6] Li, J., & Sun, J. (2026). Development of digital financial inclusion and entrepreneurial choices of rural migrant workers. *Technoeconomics & Management Research*, (2), 68–74.
[7] Yang, Z., Yue, Q., Li, Y., et al. (2026). Functional boundaries between digital financial inclusion and informal finance: A consumption-insurance perspective. *Modern Finance Research*, (2), 16–32.
[8] Li, W., Li, Y., & Zhong, Y. (2026). Mechanisms and effects of digital financial inclusion in promoting common prosperity: Evidence from double machine learning. *Journal of South China Agricultural University (Social Science Edition), 25*(1), 90–104.
[9] Xie, J., Tian, Y., & Song, M. (2025). Specialty agriculture development, rural entrepreneurial activity, and the urban-rural income gap at the county level. *Statistics & Decision, 41*(21), 48–53.
[10] Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association, 113*(523), 1228–1242.