

Digital Financial Capability and Household Wealth Portfolio Diversification in China: Evidence from a CHFS-Derived Balanced Panel

Silang Zeng*

Guangdong University of Finance, Guangzhou, China

**Corresponding author: Silang Zeng.*

Abstract

Objectives. This study examines whether within-household changes in digital financial capability are associated with changes in household wealth portfolio diversification in China. **Data and Methods.** We analyse a publicly released, unweighted balanced panel derived from the 2017 and 2019 China Household Finance Survey, comprising 18,396 household-year observations for 9,198 households. Diversification is measured by one minus the Herfindahl–Hirschman index across five asset categories, asset count, and normalised entropy. Household fixed-effects models include observed time-varying controls and a survey-year effect, with standard errors clustered by household. **Results.** The fixed-effects coefficient for digital financial capability on one-minus-HHI is 0.3476 (95% confidence interval from 0.2617 to 0.4335). A one-standard-deviation increase in capability corresponds to an approximately 0.030 increase, or 8.6% of the outcome mean. A household bootstrap with 500 replicates gives a percentile interval from 0.2478 to 0.4328. Positive associations also appear for asset count and normalised entropy, and remain under strict original-ratio and asset-count threshold checks. Component regressions associate higher capability with a lower housing share and higher productive, non-risky financial, and risky financial shares. Rural and income interactions are not statistically distinguishable from zero. **Conclusions.** Digital financial capability is positively associated with portfolio diversification within this released two-wave sample. The design does not identify a causal effect and, because survey weights and city identifiers are unavailable, does not support population-representative or subnational claims.

Keywords

household portfolio diversification; digital financial capability; household finance; China Household Finance Survey; panel data

1. Introduction

Household portfolios can remain concentrated even when a wider set of financial products is available. Such concentration ties household balance sheets to a small number of asset-specific shocks and narrows the range of saving and investment choices. Diversification is not equivalent to risky-market entry: a household

may acquire one financial product while retaining most wealth in a dominant category. Evidence from Australia illustrates systematic variation in household portfolio diversification [1]. In China, the issue is increasingly connected to digital access, payment use, financial knowledge, and the ability to navigate formal financial services.

Research on Chinese households generally associates digital finance with financial-market participation. Regional digital-finance measures and policy-led expansion are linked to risky-market entry [2–4], while internet use and mobile payments are associated with participation and portfolio choice [5–7]. Financial literacy and attention also shape how households process information and select products [8, 9]. These studies establish the importance of access and use, but participation alone does not reveal whether wealth becomes more evenly distributed across asset categories. The exposure level also matters for interpretation. A regional index describes the financial-service environment shared by households in a location, whereas household capability records whether a family can understand and use those services. Within-household changes in capability therefore answer a different question from cross-regional differences in digital infrastructure.

A related literature examines allocation beyond market entry. Household digitisation and financial inclusion have been linked to portfolio composition and selection [10, 11], and institutional or life-course conditions also influence household asset allocation [12–14]. Two gaps remain. First, binary participation and risky-asset shares do not jointly capture portfolio breadth and balance. A category count can increase when a household acquires a negligible holding, while a concentration index can improve without adding a new category. Entropy gives greater weight to the distribution across categories. Agreement across the three outcomes is therefore more informative than any measure alone. Second, cross-sectional estimates combine persistent differences between households with changes within the same household. Household fixed effects shift attention to within-household variation, although they do not remove time-varying confounding or establish causality.

We analyse a directly downloadable balanced panel derived from the 2017 and 2019 China Household Finance Survey. The main outcome is one minus the Herfindahl–Hirschman index across five asset categories; asset count and normalised entropy provide breadth and balance checks. The analysis combines household fixed effects with strict composition samples, economically meaningful holding thresholds, household bootstrap inference, and asset-component regressions. It asks whether digital financial capability and diversification move together within the same households after observed time-varying controls and a common year effect. The contribution has three parts. First, the same empirical specification is evaluated across concentration, breadth, and balance outcomes. Second, threshold and original-share checks make measurement sensitivity visible rather than treating the released ratios as exact compositions. Third, component regressions and null interaction results show what the aggregate association does and does not resolve. The estimates remain associational and apply only to the unweighted released sample.

2. Data and Empirical Strategy

2.1 Data Source and Analytic Sample

The analysis uses a publicly downloadable household-year file derived from the 2017 and 2019 waves of the China Household Finance Survey (CHFS), archived by Science Data Bank under DOI 10.57760/sciencedb.j00214.00116. The balanced panel contains 18,396 observations for 9,198 households. No survey-weight field or province/city identifier is available, so the estimates are unweighted and do not represent national or subnational populations.

The source records housing, productive assets, other non-financial assets, non-risky financial assets, and risky financial assets. Separate winsorisation leaves 44 rows with a negative component and 109 with a component above one. The median sum of the five released ratios is 0.983, with 5th and 95th percentiles of 0.604 and 1.170. For the primary construction, negative values are set to zero and the five non-negative components are renormalised. Requiring the original five-share sum to lie within 0.10, 0.05, or 0.02 of one in both waves retains 9,848, 6,890, or 3,826 observations, respectively. The central strict sample also requires every original share to lie in $[0,1]$ and contains 3,445 households. Its mean capability and one-minus-HHI are 0.061 and 0.298, compared with 0.091 and 0.379 among excluded rows. The strict sample is therefore a sensitivity check, not a preferred population.

2.2 Exposure, Outcomes, and Controls

Digital financial capability is an entropy-weighted household composite of 12 indicators spanning five literacy items, five behaviour items, and two accessibility items [15]. Its mean is 0.0800, its standard deviation is 0.0862, and its observed range is 0.0001–0.7143. The primary outcome is one minus the Herfindahl–Hirschman concentration index,

$$D_{it} = 1 - \sum_{j=1}^5 s_{ijt}^2, \quad (1)$$

where s_{ijt} is household i 's renormalised share of asset category j in wave t . Higher values indicate a less concentrated portfolio. Two alternative outcomes capture different aspects of diversification. The main asset count is the number of categories with a share above 10^{-8} ; sensitivity counts instead require shares above 1% or 5%. Normalised entropy is

$$E_{it} = - \frac{\sum_{j=1}^5 s_{ijt} \log(s_{ijt})}{\log(5)}, \quad (2)$$

with zero-share terms defined as zero.

The household fixed-effects specification includes age and age squared, health, household size, numbers of children and older members, log income, log consumption, log debt, local price level, and local economic level. The pooled specification additionally includes gender, education, marital status, social insurance, rural status, macro-region, and financial-industry employment. Household fixed effects absorb time-invariant household characteristics, and all stated time-varying controls remain identified in the within-household specification.

2.3 Estimator and Inference

The primary model is

$$D_{it} = \beta C_{it} + \gamma' X_{it} + \alpha_i + \lambda_t + \varepsilon_{it}, \quad (3)$$

where C_{it} denotes digital financial capability, X_{it} contains time-varying controls, α_i is a household fixed effect, and λ_t denotes the survey-year fixed effect. Standard errors are clustered by household. The coefficient β is interpreted as the conditional within-household association between changes in capability and changes in diversification. In this balanced two-period sample, fixed effects and first differences use identical complete-case observations and controls; under these conditions, their point estimates agree up to numerical precision.

The model comparison reports pooled ordinary least squares and household fixed effects for the three outcomes. In this balanced two-period setting, first differences use one change record per household and reproduce the fixed-effects point estimate under identical controls. They are therefore an implementation check rather than independent evidence. We also estimate the primary model without time-varying controls, without the year effect, under three strict original-share sum tolerances, and with 1% and 5% asset-count thresholds. For sampling uncertainty, we resample the 9,198 household change records with replacement in 500 household-level bootstrap replications and report percentile confidence intervals.

Asset-share regressions describe which portfolio components move with digital financial capability; they do not identify mediation. Rural-status and log-income interactions are secondary tests, and their null results are retained rather than converted into subgroup claims.

3. Results

3.1 Sample and Primary Association

The balanced panel contains 18,396 household-year observations for 9,198 households. Mean asset count, one-minus-HHI, and normalised entropy are 3.520, 0.349, and 0.391, respectively; the strict sample contains 6,890 observations from 3,445 households.

The household fixed-effects estimate is the main result. A one-unit capability change is associated with a 0.3476 change in one-minus-HHI after controls and the year effect. At the observed capability standard deviation of 0.0862, this corresponds to an approximately 0.030 increase, or 8.6% of the 0.3488 outcome mean.

These are associational quantities, not causal effects. The household-clustered 95% CI for the one-unit coefficient is 0.2617–0.4335 ($p < 0.001$). Table 1 shows that pooled OLS gives a larger coefficient, 0.4490, indicating that between-household differences contribute to the pooled relationship. First differencing reproduces the fixed-effects coefficient only because it is algebraically equivalent under the balanced complete-case conditions described in Section 2.

Table 1: Digital financial capability and portfolio diversification

Estimator and outcome	Coefficient	SE	95% CI	N
Pooled OLS: one-minus-HHI	0.4490	0.0232	[0.4035, 0.4945]	18,396
Household FE: one-minus-HHI	0.3476	0.0438	[0.2617, 0.4335]	18,396
Household FE: asset count	0.9717	0.1501	[0.6776, 1.2658]	18,396
Household FE: normalised entropy	0.3794	0.0450	[0.2912, 0.4677]	18,396
Household FE: one-minus-HHI, strict sample	0.2925	0.0816	[0.1326, 0.4525]	6,890

Notes: Entries are coefficients on digital financial capability. Standard errors are clustered by household. Full-sample regressions contain 9,198 households; the strict sample contains 3,445.

3.2 Robustness and Portfolio Composition

The fixed-effects association remains positive for asset count and normalised entropy (Table 1). Their coefficients are 0.9717 and 0.3794, with 95% CIs of 0.6776–1.2658 and 0.2912–0.4677. Requiring a category to exceed 1% or 5% of the portfolio gives count coefficients of 1.5774 and 1.3270; the corresponding intervals are 1.1950–1.9598 and 0.9294–1.7246. Thus, the breadth result is not driven by numerical traces near zero.

The central strict-sample coefficient is 0.2925, with a 95% CI of 0.1326–0.4525 (Table 1). Figure 1 places the primary estimate beside pooled, reduced-control, and strict-sample variants. The estimates without time-varying controls and without the year effect are 0.3857 and 0.3469. Tightening or relaxing the original-share tolerance changes the magnitude and precision because it changes the sample, but none of the displayed estimates changes sign.

The bootstrap mean is 0.3455, with a percentile 95% interval of 0.2478–0.4328. This interval is close to the household-clustered interval and remains wholly positive, although the bootstrap is a sensitivity analysis rather than a new identification strategy. Table 2 reports threshold and portfolio-component results without repeating the specifications shown in Figure 1.

The component regressions associate higher capability with a lower housing share (−0.2686) and higher productive, non-risky financial, and risky financial shares (0.0925, 0.1072, and 0.1123). The other non-financial estimate is −0.0434, and its interval includes zero. Because the five shares sum to one, these coefficients describe a joint reallocation pattern rather than mediation or a sequence of independent effects.

The rural interaction is −0.0849 (95% CI, −0.2804 to 0.1106), whereas the log-income interaction is 0.0290 (95% CI, −0.0135 to 0.0715). Neither interaction is statistically distinguishable from zero. The evidence therefore does not support rural–urban or income heterogeneity claims in this sample, although the imprecision does not establish homogeneous relationships.

Figure 1: One-minus-HHI coefficients across estimator and sample variants. Horizontal lines are 95% confidence intervals; pooled and panel estimators use household-clustered standard errors

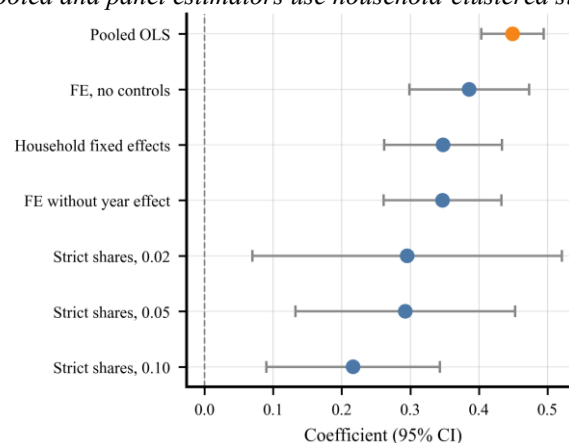


Table 2: Threshold and portfolio-component estimates

Outcome or term	Coefficient	95% CI	N
Panel A. Asset-count thresholds			
Asset count, share above 1%	1.5774	[1.1950, 1.9598]	18,396
Asset count, share above 5%	1.3270	[0.9294, 1.7246]	18,396
Panel B. Portfolio components			
Housing share	-0.2686	[-0.3866, -0.1507]	18,396
Productive-asset share	0.0925	[0.0091, 0.1759]	18,396
Other non-financial share	-0.0434	[-0.1155, 0.0287]	18,396
Non-risky financial share	0.1072	[0.0451, 0.1693]	18,396
Risky financial share	0.1123	[0.0603, 0.1643]	18,396

4. Discussion

The positive association across concentration, category count, and entropy suggests that digital financial capability is related to portfolio structure rather than only market entry. The component estimates are consistent with movement away from housing concentration toward productive and financial categories, although the design cannot identify the mechanism. Capability may reduce information and transaction barriers, but income expectations, changing liabilities, or unmeasured financial advice could also affect both capability and portfolio choice. This interpretation extends evidence linking digital finance, mobile payment, and internet access to market participation [2, 5, 6] and is consistent with studies of diversification and allocation efficiency [7, 11, 13]. The distinction matters because risky-market entry records access or willingness, whereas diversification concerns the distribution of the whole balance sheet. Following the same households across two waves reduces dependence on persistent between-household differences, while agreement across three outcomes reduces dependence on one measurement choice.

The evidence remains associational. Time-varying omitted factors and reverse association may persist after household fixed effects, and two waves cannot support a dynamic or causal interpretation. A policy discontinuity, defensible instrument, or comparable source of exogenous capability variation would be required for a causal claim. The released file contains neither survey weights nor province or city identifiers, so the results are not population representative and cannot identify Greater Bay Area patterns. The strict sample is also selected on observed capability, income, and portfolio concentration, which explains its use as a sensitivity check rather than a preferred estimate. Finally, the null rural and income interactions indicate limited precision rather than homogeneous relationships. Longer weighted panels with documented subnational identifiers are required to test causal and distributional questions.

5. Conclusion

Using a public 2017–2019 CHFS-derived balanced panel, we find that within-household increases in digital financial capability are positively associated with broader and more balanced wealth portfolios. Agreement across concentration, asset-count, entropy, threshold, and strict-sample checks indicates that the result is not tied to one diversification definition or data-construction rule. Component estimates are consistent with movement away from housing concentration, while rural and income differences remain unresolved. The two-wave fixed-effects design does not establish causality or population effects. Longer weighted panels with subnational identifiers are needed to test causal and distributional questions.

References

- [1] Worthington, A. C., “Household asset portfolio diversification: Evidence from the household, income and labour dynamics in australia survey,” *SSRN Electronic Journal*, 2009, <https://doi.org/10.2139/SSRN.1421567>.
- [2] Ye, Y., Yong-jian, P., and Xiong, A., “The impact of digital finance on household participation in risky financial markets: Evidence-based study from china,” *PLOS ONE*, 17, (4), e0265606, 2022, <https://doi.org/10.1371/journal.pone.0265606>.

- [3] Xing, L., Han, D., and Xie, H., “Government digital policy breaks the mystery of limited participation in china’s home finance market,” *Scientific Reports*, 13, (1), 2023, <https://doi.org/10.1038/s41598-023-47372-6>.
- [4] Zhang, Y., Xu, J., and Chen, D., “Fintech and household participation in risky financial markets: Evidence from china,” *Financial Innovation*, 12, (1), 2026, <https://doi.org/10.1186/s40854-025-00845-1>.
- [5] Peng, G., Peng, X., and Zhu, L., “Internet use, household financial market participation and portfolio choice: Evidence from china,” *China Finance Review International*, 16, (2), 373–395, 2024, <https://doi.org/10.1108/cfri-10-2023-0281>.
- [6] Xu, Y., “The impact study of mobile payment on household asset allocation: Evidence from CHFS data,” *Journal of Applied Economics and Policy Studies*, 16, (1), 30–40, 2025, <https://doi.org/10.54254/2977-5701/2025.20684>.
- [7] Qiu, W., “Does mobile payment adoption increase household portfolio diversification? Evidence from china,” *Finance Research Letters*, 75, 106911, 2025, <https://doi.org/10.1016/j.frl.2025.106911>.
- [8] Xu, S., Yang, Z., Ali, S. T., Li, Y., and Cui, J., “Does financial literacy affect household financial behavior? The role of limited attention,” *Frontiers in Psychology*, 13, 2022, <https://doi.org/10.3389/fpsyg.2022.906153>.
- [9] Wang, S., Cao, P., and Huang, S., “Household financial literacy and relative poverty: An analysis of the psychology of poverty and market participation,” *Frontiers in Psychology*, 13, 2022, <https://doi.org/10.3389/fpsyg.2022.898486>.
- [10] Zhang, R., and Hu, S., “Digital prosperity or digital worry? Family digitization and financial asset allocation,” *International Journal of Finance & Economics*, 2026, <https://doi.org/10.1002/ijfe.70237>.
- [11] Bian, Y., Wang, X., and Zhang, Q., *How does china’s household portfolio selection vary with financial inclusion?*, 2023, <https://arxiv.org/abs/2311.01206>.
- [12] Zheng, C., Chen, Y., and Huang, Y., “Can breaking the rigid payment lead to better household financial asset allocation? Evidence from china’s financial regulation,” *Applied Economics Letters*, 33, (12), 2046–2050, 2025, <https://doi.org/10.1080/13504851.2025.2462710>.
- [13] Shen, L., and Yang, H. T., “Can the development of digital finance improve the financial asset allocation efficiency of middle class households?” *SHS Web of Conferences*, 163, 01030, 2023, <https://doi.org/10.1051/shsconf/202316301030>.
- [14] Xu, S., and Yang, Z., “Early-life famine experience and households’ diversity of financial asset portfolio: Evidence from china,” *Social Science Quarterly*, 106, (2), 2025, <https://doi.org/10.1111/ssqu.13493>.
- [15] Liu, C., Zhang, Y., and Li, G., “Digital financial capability and household wealth accumulation—evidence from the china household finance survey,” *China Journal of Econometrics*, 5, (2), 442–462, 2025, <https://doi.org/10.12012/CJoE2024-0264>.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Data and Code Availability

The household-level source file used in this study was obtained from the Science Data Bank archive associated with Liu et al. [15] (dataset DOI: 10.57760/sciencedb.j00214.00116). This DOI identifies the source-data archive rather than the present manuscript. In accordance with the archive's CC BY-NC-ND 4.0 licence, we do not redistribute cleaned or transformed respondent-level records. Replication code, variable-construction documentation, checksums, and aggregate results can be provided upon reasonable request, subject to the source-data licence.

Open Access

This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

