Understanding Socio-Economic Drivers of Mobile Banking Adoption

In this section, we employed a rigorous analytical approach to investigate the determinants of mobile banking adoption, focusing on socio-economic and demographic factors. Initially, we conducted fixed and random effects regressions and utilized a simple Hausman test to evaluate the consistency of coefficients between the fixed and random effects models to determine the most appropriate modelling technique for our panel data. The significant outcome of this test ($p = 0.0000 < 0.05$) indicates that the assumptions underlying the RE model were violated, and so an FE approach was more suitable.

To ensure the reliability of our regression analysis, we conducted Levin-Lin-Chu (LLC) unit root tests on all variables. Stationarity is a critical assumption in time series analysis, as non-stationary variables can lead to spurious regression results and unreliable inference. The LLC unit root test examines whether a variable contains a unit root, indicating non-stationarity, or if it is stationary over time. We found that the variable "Urbanpopulationoftotalpop" exhibited non-stationarity initially, prompting us to apply a logarithmic transformation to stabilise its variance and achieve stationarity. Additionally, for variables where stationarity could not be determined due to missing data, such as "IndividualsusingtheInternet" and "Schoolenrollmentprimarygr," we proceeded with the analysis, acknowledging the potential limitations. Furthermore, the variable "gender_ratiomaletofemale" was found to be non-stationary, despite applying a logarithmic transform. A first difference transformation solved the issue but eliminated the data for 2015 for every country, which introduced systematic missingness and which in turn could introduce bias. Thus, we decided to leave the variable unchanged.

Table 5 Robust fixed-effects estimation

VARIABLES	Numregisteredaccounts
L_Urbanpopulationoftotalpop	6,563***
	(2,120)
gender_ratiomaletofemale	-6,193
	(12,409)
GDPpercapitaconstant2015US	0.174
	(0.180)

Youthpopulationages1524	-159.7
	(127.0)
Populationages65andabove	-283.0**
	(120.9)
IndividualsusingtheInternet	22.70***
	(8.055)
Employmenttopopulationratio	16.18
	(26.86)
Schoolenrollmentprimarygr	-29.25**
	(10.75)
Constant	-11,520
	(16,548)
Observations	154
Number of country_id	25
R-squared	0.722
F(8,24), Prob > F	10.6, 0.0000
Modified Wald Test: chi2 (25)	1561.81
Prob>chi2	0.0000

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5 gives the results for our fixed effects model with robust standard errors to enhance the robustness of our analysis. This approach allows for the clustering of standard errors at the country level, thereby accommodating the potential correlation of errors within each country while mitigating the effects of heteroskedasticity. As expected, urban population and individuals using the internet were positively significant with the dependent variable at the 5% level, and population aged 65+ were negatively significant at the 10% level. There is also evidence to suggest that primary school enrolment has a negative relationship with the dependent variable, although with considerably less magnitude than the population aged 65+ variable. Further investigation may be warranted to better understand this relationship as its significance might be context-specific or influenced by other uncaptured variables. Gender ratio, GDP per capita, youth population and the employment-to-population ratio did not yield significant results.

The within-group R-squared (0.7219) indicates a significant proportion of the dependent variable's variance is explained by the model, whilst the presence of significant effects among the independent variables is confirmed by the F-test (Prob < F = 0.0000). However, a Wald test for groupwise heteroskedasticity revealed high heteroskedasticity in the model (Prob>chi2 = 0.0000), indicating potential issues with the model's robustness. In Section 6, we will employ the Generalized Method of Moments (GMM) framework, which offers the flexibility to capture the heterogeneity observed in the model.

Socio-economic Impacts: Generalised Method of Moments (GMM) Analysis

1. *Methodology and Model*

In their study on financial development and economic growth in China, (Zhang et al., 2012) employed generalized method of moments (GMM) estimators (Hansen, 1982) to address endogeneity, serial correlation and heteroskedasticity. The GMM panel estimators enabled Zhang et al. to control for time-fixed effects and city-specific effects while using appropriate lags of independent variables as instrumental variables. Building on this, our implementation of GMM is motivated by the heterogeneity issues indicated primarily by our modified Wald test results.

There are two main types of GMM commonly used in panel data analysis: first difference GMM and system GMM.

First-difference GMM accounts for unobserved heterogeneity and endogeneity by taking first differences of the variables and using these transformed variables to estimate the model parameters. System GMM extends the first-difference approach by incorporating lagged observations of the variables as instruments. This method allows for more efficient estimation by exploiting both the within-group variation (through first differences) and the between-group variation (through lagged levels) in panel data.

After discovering that the gender_ratio variable exhibited non-stationarity despite attempted transformations (first difference transformation resulted in systematic missingness) and considering its insignificance in the fixed effects regression results, we have opted to omit it from the GMM model. Consider the following regression equation:

$Y_{it} = \alpha Y_{i,t-1} + \beta X_{it} + \eta_i + \lambda_t + \epsilon_{it}$ (3)

Country-specific effects η_i introduce inconsistency in within-group estimators due to correlation with the lagged dependent variable $Y_{i,t-1}$. To eliminate country-specific effects, the first difference of the equation is taken:

$\Delta Y_{it} = \alpha \Delta Y_{i,t-1} + \beta \Delta X_{it} + \Delta \lambda_t + \Delta \epsilon_{it}$ (4)

where Y_{it} represents the first difference of the dependent variable Numregisteredaccounts, and ΔX_{it} denotes the first difference of the lagged values of our independent variables.

We believe that our annual data will exhibit a slower temporal dynamic compared to a more high-frequency data, therefore, we will use the first lag differences of our independent variables in Eq. (4) as our additional instruments.

2. *GMM Regression Results*

Table 7 GMM Results

	(1)
VARIABLES	Numregisteredaccounts
L.Numregisteredaccounts	1.210***
	(0.0492)
L_Urbanpopulationoftotalpop	83.79**
	(33.40)
GDPpercapitaconstant2015US	-0.0262*
	(0.0156)
Youthpopulationages1524	-32.85
	(21.67)
Populationages65andabove	-7.747
	(12.34)
IndividualsusingtheInternet	-0.214
	(1.625)

Employmenttopopulationratio	-1.139
	(1.717)
Schoolenrollmentprimarygr	-0.729
	(1.002)
Constant	575.4
	(470.5)
Observations	130
Number of country_id	25

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The dynamic panel-data estimation using the one-step system GMM approach yielded several coefficient estimates for the explanatory variables. Notably, the coefficient for the lagged dependent variable (Numregisteredaccounts) is statistically significant at the 1% level, indicating a strong positive relationship between the number of registered mobile money accounts in the current year and the previous year. Among the control variables, the coefficient for the urban population as a percentage of the total population (L_Urbanpopulationoftotalpop) is statistically significant at the 5% level, suggesting that urbanization might positively influence the adoption of mobile money services. Also, GDPpercapitaconstant2015US is significant at the 10% level. However, other variables such as youth population, population aged 65 and above, individuals using the Internet, employment-to-population ratio, and primary school enrolment rate do not exhibit statistically significant associations with the number of registered mobile money accounts.

3. *Diagnostic Tests*

Instruments for both the first differences and levels equations were carefully selected to address endogeneity issues and ensure the consistency of the GMM estimators. However, diagnostic tests revealed some potential concerns.

Arellano-Bond test for AR(1) in first differences:	z = -2.25 Pr > z = 0.024

Arellano-Bond test for AR(2) in first differences:	z = -0.68 Pr > z = 0.499

The AR(1) test indicated first-order autocorrelation in the first differences of the dependent variable, suggesting a need for further investigation into the temporal dependencies within the data. Conversely, the AR(2) test showed no evidence of second-order autocorrelation. The Sargan test suggested a violation of overidentifying restrictions, indicating potential misspecification of the model. However, the Hansen test indicated that the instruments used in the GMM estimation were valid, supporting the consistency of the estimation method.

Sargan test of overid. restrictions:	chi2(22) = 38.10 Prob > chi2 = 0.018
(Not robust, but not weakened by many instruments.)	
Hansen test of overid. restrictions: chi2(16)	chi2(22) = 16.81 Prob > chi2 = 0.774
(Robust, but weakened by many instruments.)	

Furthermore, difference-in-Hansen tests confirmed the exogeneity of instrument subsets, reinforcing the validity of the GMM estimation. These results provide valuable insights into the determinants of mobile money account registration across countries, highlighting the importance of both economic and demographic factors. However, caution is warranted in interpreting the results, particularly regarding the presence of autocorrelation and potential model misspecification, which may impact the reliability of the estimated coefficients and their implications for policy and practice.

Difference-in-Hansen tests of exogeneity of instrument subsets:	
GMM instruments for levels	
Hansen test excluding group:	chi2(16) = 16.29 Prob > chi2 = 0.433
Difference (null H = exogenous):	chi2(6) = 0.52 Prob > chi2 = 0.988
iv(L_Urbanpopulationoftotalpop GDPpercapitaconstant2015US Youthpopulationages1524 Populationages65andabove IndividualsusingtheInternet Employmenttopopulationratio Schoolenrollmentprimarygr)	

Hansen test excluding group:	chi2(15) = 16.56 Prob > chi2 = 0.346
Difference (null H = exogenous):	chi2(7) = 0.25 Prob > chi2 = 1.000

Plotting Residuals

Visually, heteroskedasticity can be assessed by examining the relationship between the residuals and the fitted values of our model. If the plot indicates that the residuals tend to increase or decrease systematically as the fitted values change, it suggests the presence of meaningful variance. However, upon inspecting the graph, we observe that the residual distribution lacks a discernible pattern, indicating the absence of significant heteroskedasticity. While there might be some underlying heteroscedasticity in the residuals, employing Generalized Method of Moments (GMM) techniques along with robust standard errors ensures the generation of reliable estimates even in the presence of such heteroskedasticity.

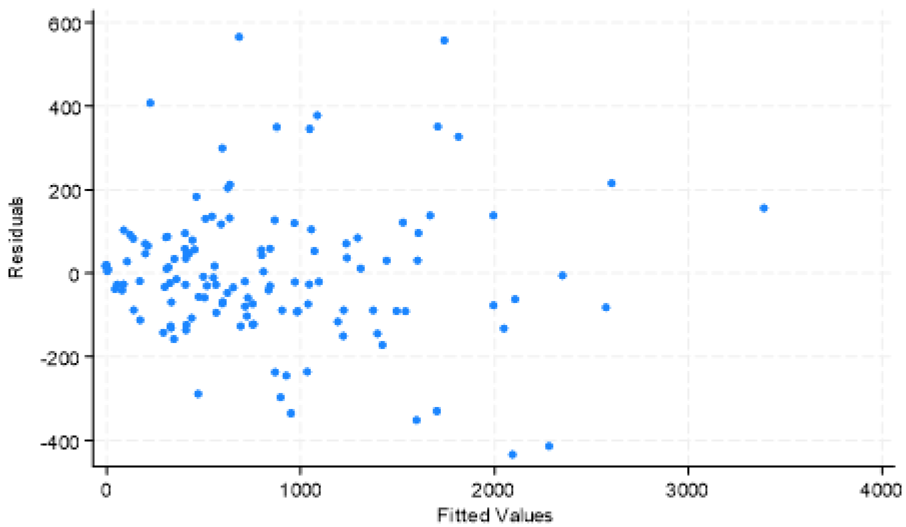

Figure 2 Plot of Residuals

Conclusions

The fixed effects (FE) model in our study revealed significant influences of variables such as urbanisation rates on the uptake of mobile banking. However, the GMM model, addressing endogeneity concerns, highlighted fewer significant factors, emphasizing the importance of accounting for dynamic relationships.

Our findings underscore the multifaceted nature of mobile banking adoption, influenced by a myriad of economic, demographic, and external factors. While each methodology provided valuable insights, their integration revealed the complexity inherent in understanding financial behaviour. Moving forward, further research could refine these models, exploring additional variables and employing more sophisticated econometric techniques to enhance predictive accuracy and policy relevance. Such endeavours hold promise for advancing financial inclusion efforts and fostering resilience in an ever-evolving digital landscape.