Three Essays in Applied Economics

by

Kumuditha Dilanka Hikkaduwa Epa Liyanage

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
May 7, 2022

Keywords: Land use, Irrigation, Financial inclusion, Unbanked households,
Rural households, Prepaid cards

Approved by

Denis A. Nadolnyak, Chair, Professor, Agricultural Economics
Valentina Hartarska, Co-Chair, Alumni Professor, Agricultural Economics & Finance
Wenying Li, Member, Assistant Professor, Agricultural Economics
Yaoqi Zhang, Member, Alumni Professor, Forest Economics and Management
Jitka Hilliard, University Reader, Professor, Finance

Abstract

This dissertation includes three empirical essays as three chapters selecting different research questions in Land Economics and Financial and Development Economics. I contribute to the literature by extending and implementing methods of dynamic panel regression, nonlinear and nonparametric model estimation, and estimating causal effects to address the issues of heterogeneity, endogeneity, and confounding factors.

Chapter 1 determines the association between agricultural land-use change, cropping patterns, and land and irrigation determinants in Alabama, using United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS), Forest inventory analysis (FIA), USDA soil survey, and United States Geological Survey (USGS) aggregate panel data for the period from 1992 to 2017 on the county level. This study integrates the dynamic panel regression model using Arellano-Bond (AB) estimator comparing other models such as fixed effects (FE), seemingly unrelated regression (SUR), and ordinary least squares (OLS) and finds that major land attributes such as Agriculture and forest revenues and expenses show negative and positive effects on land use shares and irrigated land shares show positive effects for cropland, whereas more detailed irrigation variables are insignificant with land use shares. Overall, crop shares are less associated with the explanatory variables as the distribution of crop acreage within the cropland in the state with different local climates and soils and crop shares of cotton and corn are more associated with land determinants compared to peanut and soybean shares. Principal component analysis confirms the initial assumption that land-specific economic and social attributes are associated with land-use change and cropping decisions in Alabama.

Chapter 2 contains an impact analysis of financial inclusion of unbanked rural households after the Dodd-Frank act of 2010 using the Quantile treatment effect (QTE) with the nonlinear

Changes-in-changes (CIC) model utilizing 2009, 2011, and 2019 household data from the National household survey conducted by FDIC. The discrete CIC model is developed using the DID approach and estimated with quantile treatment effects for the unbanked rural and urban households' variables. Results indicate that the Dodd-Frank act is associated with a higher likelihood of opening a bank account by the unbanked rural households' in both the short and long term with an expected smaller long-term magnitude. Relative to urban unbanked households, rural unbanked households are more likely to use AFS for transaction purposes in short term, indicating a strong potential effect related to substituting transactions via banks with transactions via AFS. However, relative to urban households, rural households are more likely to have used more AFS for credit purposes, in the long run in 2019, which may be related to the resulting closures of banking infrastructure from 2009 to 2019.

Chapter 3 determines the factors associated with prepaid debit card usage of unbanked households' impact on financial inclusion entering formal banking services and additionally, investigate the use of alternative financial services as a substitute for traditional banking in the US. The results show that unbanked prepaid debit cards users are less likely to open a bank account. In further investigation, adding AFS credit and transaction services as outcome variables, results show that unbanked prepaid card users are more likely to use alternative financial services. These findings suggest that unbanked households who use prepaid cards have a tendency to use more alternative financial services compared to the more valuable and typically cheaper banking services offered by traditional banks for their financial needs in the household system in the US. Overall, the study confirms that unbanked households who use prepaid debit cards are open to using financial instruments and services that meet their needs outside of the formal financial system with a propensity to use alternative financial services.

Keywords: Land use, Irrigation, Financial inclusion, Unbanked households,
Rural households, Prepaid debit cards

Acknowledgments

First and foremost, I would like to thank and express my heartfelt gratitude to Dr. Denis Nadolnyak, chair of my committee, and Dr. Valentina Hartarska, co-chair for their continuous guidance, invaluable advice, endless support, and being my mentors throughout my doctoral journey. Their immense knowledge and ample experience encouraged me throughout this time. Having their continuous guidance during this period was a silent blessing that helped me grow into the person I am today. Without them, this would have not been possible.

Also, I would like to thank my committee members for their time and valuable opinions. In particular, I would like to thank Dr. Wenying Li for his kind support and guidance throughout my last year of the program also many thanks and appreciation to Dr. Yaoqi Zhang for his suggestions and patience with me throughout the research. Also, I would like to thank Dr. Jitka Hilliard, my University Reader who accepted my request to serve on my committee in a short time to get me to the end of this journey.

I would also like to thank my friends for their support and encouragement during my research work. Special thanks to the department head Dr. Joshua Duke and administrative staff members, Ms. Ashley S. Pangle, Ms. Ann M. Gulatte, and Ms. Connie Pope who work in the Department of Agricultural Economics & Rural Sociology for all the help they have provided during my journey.

Finally, I would like to express my gratitude to my dear wife, Prabha Liyanapathiranage, and my son Sayul. Without their tremendous understanding and encouragement in the past few years, it would be impossible for me to complete my journey.

Last but not least, I would like to thank my parents, my in-laws, my brother, and my sister for their support and unconditional love.

Table of Contents

List of Tables

List of Figures

1. CHAPTER 1

Impacts of irrigation and land determinants on land-use change and cropping patterns in Alabama

Introduction

Land use plays a major role in agricultural production, and it is important for social, economic and urban development. Land use models have been studied more than three decades and new models are developing with many factors that influence on land use change. Land use decisions in Agriculture have impacted production and environmental outcomes in short and long term and more analysis of economic and land quality characteristics is needed to better understand land use decisions and for policy reference (Claassen and Tegene 1999). When agricultural policies have direct and indirect effects on land use decisions, land use changes and cropping patterns can be significantly impacted by factors such as irrigation use and other land characteristics (Lubowski, 2006). Physical factors such as land quality and landscape, socio-economic factors such as population growth, market conditions, income and government policies are main contributing factors of land use change that are supported by past studies (Majumdar, 2009). In the United States, major land use changes have taken place during the past 25 years and the extent to which land will be converted from other land into crop lands has become a key issue facing US agriculture (Wu, 2008).

The southeastern US has undergone a significant land-use change since 1950 and continues to experience higher demand for forestry and agriculture with forest land conversion, population growth, urban development and agricultural land use changes (Napton et al., 2010, Sohl and Sayler, 2008). Southeast US has been a region with abundant surface and ground water resources and was able to meet the water demand, even though the region was less profitable with low soil quality and lack of irrigation compared to other regions. In the second half of the twentieth century, the region experienced significant land use conversions with the expansion of irrigated land, which increased the pressure on water demand of the region (Mullen et al., 2009). The expanded irrigated acreage also reflected cropping pattern changes, supply of abundant aquifer water sources, and sensitivity of producers to changing precipitation levels in the region.(USDA-ERS 2021).

The objective of this paper is to investigate the effects of irrigation, land quality and other factors on land use and cropping patterns in Alabama. Alabama has comparatively rich water supply related to other states in the region, but its irrigated crop land share is very low (1.3% of the national average). According to the 2017 United States Department of Agriculture (USDA) Census data, the cropland share is approximately 8.5% and forestland share is 70% out of total regional land showing the highest share in the southeast region. Alabama has experienced significant land use conversion of farmland to forest land or urban land. Conversion of forest land to urban land has led to a decrease of cropland and an increase of forest and urban lands over the past years (Nagubadi and Zhang, 2005). When Agriculture is the largest consumer of water, efficient irrigation has become more important for the Southeast because of increasing irrigation technology cost. Although Alabama receives average annual precipitation of over 1200 mm, agricultural regions of the state experience crop yield reduction due to the seasonal drought and uncertainty of rainfall and use of irrigation in Alabama has increased in the last two decades (Magliocca et al., 2020). The role of irrigation remains less clear and estimating how irrigation attributes and other land characteristics impact on land use conversions and crop patterns of interest is important to further explore the long term relationships between land, irrigation attributes, and land use and cropping patterns in Alabama.

Literature Review

Land use change can be categorized into two major types (Davis et al., 2019). The first type is a land cover change coupled with the expansion or reduction in the land area for different broad land use purposes such as pastureland, cropland, forest and urban land. The second type is a management type of change on existing agricultural land cover such as changes in crop type, irrigation, fertilizer use, harvesting practices etc. Without changes in the extent of different land covers, land use change can take place anywhere. Generally, land use studies explore the relationship between land use changes and explanatory variables derived from different land uses, or proxies. These variables can be spatial, economic and demographic factors such as subsidies, input and output prices, and land and climatic factors like slope, quality, and precipitation etc. (Chakir and Lungarska, 2015).

Land use determinants usually operate as proxies for land use rents that cannot be observed directly, and these proxies are changing across many studies. Most of the proxies frequently used for agriculture are producer's revenue, agricultural land price, input or output prices, yields, land quality, and government payments (Stavins & Jaffe, 1990; Wu & Segerson, 1995; Plantinga, 1996; Plantinga & Ahn, 2002). Also, previous studies have investigated production and socio-economic factors in explaining land use changes (Lichtenberg 1989, Parks and Murray 1994, Wu and Brorsen 1995, Plantinga 1996, Hardie and Parks 1997, Plantinga and Miller 1999, Ahn et al. 2000 and Lubowski et al. 2006). Land use change is also impacted by market demand and government policies (Veldkamp & Lambin, 2001). Relationships between land quality and the effect of technology or policy on land use shares' allocation have been observed in some studies (Lichtenberg 1989 and Stavins and Jaffe, 1990).

In these land use studies, researchers have constructed empirical land use models using two approaches known as aggregated method and spatial method (Majumdar et al., 2009). The aggregate approach uses socio-economic and land variables to examine the effect of land use change in different land types such as agriculture, urban and forestry (Alig 1986, Lichtenberg 1989, Parks and Murray 1994, Wu and Segerson 1995, Hardie and Parks 1997, Miller and Plantinga 1999, Ahn et al. 2000). This approach has been used in the majority of econometric land use models and collecting aggregate data can be done with low cost. Spatial approach use the basis of pixels, parcels or sample points to explore the land use or land cover change (Bockstael 1996, Claassen and Tegene 1999, Kline et al. 1999, Lubowski 2002, Polyakov and Zhang 2008).

Aggregate land use models are mostly suitable to analyze land use policies as they use historical data to capture the decisions made by private landowners who are dealing with returns that convert to alternative uses of land (Bell et al., 2017). Aggregate data of land use at the county-level, are collected by several government agencies and these data are usually available to researchers and can be retrieved according to their estimation and research purposes. Selecting appropriate variables in land use models is important for estimating the relationship between land use shares and land use determinants in common aggregate data models (Lichtenberg 1989; Parks and Murray 1994; Wu and Segerson 1995; Hardie and Parks 1997; Plantinga et al. 1999). When county data is utilized in the model, land use shares defined as percent of land uses of total county area (Bell et al., 2017).

Revenue is an important variable that is used in land use model in most of land use studies and actual revenue that comes from either agriculture or forestland use is the reason for landowner’s tendency to convert to alternative land uses or continue with same land use. (Adjei, 2020). For example, Alig (1986) used crop, timber and beef income variables to examine the major land use changes in the southeast United States and Hardie et al. (2000)

applied agricultural income and timber rent as proxies for actual revenue to study land use changes. When the returns of agricultural and forestry land uses depend on physical land characteristics such as slope, soil type, and water-holding capacity, many land use studies have extended land use models to include these factors (Lichtenberg 1989, Plantinga 1996, Hardie and Parks 1997 and Miller and Plantinga 1999).

Lichtenberg (1989) concentrated on the land quality impact on land use, crop choice, and irrigation development with technological change. He collected county level data of land use shares for seven major crops in western Nebraska including variables such as crop prices, irrigation technology cost, and land quality measuring average water capacity of soil. As the land quality had a significant impact on cropland allocation, technology was important applying for low quality land or else irrigation was sensitive to tax policies. Hardie and Parks (1997) explored a land use model using county level observations of irrigated farmland, other farmland, and forestland shares in the southeastern US. This area based model was built to examine the effect of land quality including independent variables such as crop revenue and cost, timber prices and costs, land class measuring land quality, and sociodemographic factors (average age of landowners, population density, per capita income and dummies for regions) maximizing the land rent.

Also, several studies have considered how land determinants such as farm and forest income, population density, agriculture and timber prices, land quality, and government programs impact land use changes adopting regression models such as modified linear, logistic and seemingly unrelated regression for the analysis (Alig, 1986, Park and Murray 1994, Plantinga and Miller 1999, Ahn et al., 2000, Nagubadi and Zhang, 2005 and Polyakov and Zhang, 2008). Alig (1986) and Majumdar et al., (2007) also indicated that population density, income, government programs, and regional dummies have effect on land use changes.

Plantinga (1996) and Miller and Plantinga (1999) estimated more basic shares models that estimate revenue of land-use shares by land quality class. They estimated the forest land shares for each land quality class instead of forest share for the entire county. Plantinga and Miller (1999) showed that population density had a positive effect on urban land relative to other land types. In addition, agricultural revenue had positive effect on agricultural land shares in Maine, South Carolina, and Wisconsin. Also, Ahn et al. (2000) showed that agricultural rent have a propensity to increase the agricultural land share relative to forest land share and further showed that higher average quality land be likely to have more agricultural land relative to forestland. Furthermore, Nagubadi and Zhang (2005) revealed that timber prices, population density, agricultural prices, and land quality variables affected land use changes and showed that higher forest return and population density increased timberland shares while lands with good quality decreased timberland shares. These results showed that farmlands had changing trends by decreasing farmland acreage and increasing timberland and urban land acreages in different proportions over the years. Lubowski et al. (2008) focused on land use change driving factors in the US between 1982 and 1997 and adopted the net returns as the drivers of land-use change. Factors from both supply and demand sides are also included in the model and private land-use decisions impacted on land quality, economic returns, and policies. Mu, et al. (2017) investigated farmers' adaptation behavior difference under future climate when incorporating future socio-economic situations and found that crop net returns are sensitive to climate variables than livestock net returns which are both increasing and show impacts on agricultural land uses are considerably different.

The Southeast region is one of the most diverse agricultural production regions in the US and Agriculture plays a major role in the region's economy. Unpredictable seasonal and spatial climate conditions with temperature and rainfall distributions, soil types, and access to irrigation are reasons to establish an intensive and diverse agricultural sector in the region (Asseng, 2013). Compared to the western US states, irrigation has received little attention in the southeastern states earlier. Since the mid-1990s, irrigated crop area has extended significantly across southeastern US according to USDA-NASS data (Lambert et al., 2021). Improved crop yields and beneficial commodity prices, low energy costs, more frequent drought conditions and abundant groundwater sources availability have caused the expansion of irrigated areas in humid regions in southeast US (Schaible and Aillery, 2012, Lambert et al., 2021). Earlier in the 20th century, agriculture in the southeast region less profitable compared to other regions due to lack of irrigation and low soil quality and region has faced farmland into forestland conversion in the latter half of the 20th century (Zhao et al., 2013). Also, Lin et al., 2000 found that generally row crop acreage is less stable because of the significant variation in local climates, ground and surface water accessibility, size and type of farms in southeast US.

Alabama’s annual precipitation is more than 1200mm and due to the large amount of rainfall, crop production is mostly dependent on rain-fed agriculture (Srivastava et al., 2010). However, the use of irrigation has increased because most precipitation occurs in the non-growing season, uncertainty associated with rainfall distribution and the severe droughts during the growing season can cause losses in crop yield and production (Srivastava et al., 2010, Lamb et al., 2011) and Alabama mainly rely on irrigation during the growing season. Even though irrigation adoption in Alabama is low compared to neighboring states, it is important to study how irrigation factors effect on land use change and cropping patterns in the state for past 2-3 decades. Considering the influence of irrigation is important for understanding land use and crop acreage allocation while determining other socio economic factors contributing to the land use changes.

This paper contributes to this literature by integrating the dynamic panel regression model (Arellano-Bond estimator) to estimate the impact of irrigation, land quality and other socioeconomic determinants on land use change and cropping patterns and compares the results to those of Ordinary Least Square (OLS), Fixed effects (FE), and Seemingly Unrelated Regression (SUR) models. This paper proceeds as follows. The next section discusses the econometric model and data used for the analysis in this paper. The following section analyzes the results of the estimations. The last section concludes and discusses possible extensions of this paper.

Conceptual Framework

Theoretically, land will be committed to the land use that yields the greatest returns to the land. Ricardian land rent approach generally assumes that land use observed in a period is consistent with a landowner who is trying to maximize returns in a risk neutral situation (Kim et al., 2018). These returns to land are often expressed in terms of land rent as the portion of total returns that adds to the land net total costs (Barlowe, 1978). The early contributions of Ricardo and von Thunen's land rent theories have been helped to build the modern land use theory. When total acreage of each region is constant, the purpose of land allocation is maximizing the total profit from all types of land. Associated hypotheses of land use share determinants construct in the form of land use equations with the shares of the major land uses in the total land and crop-specific land shares of total cropland as functions of economic, demographic, irrigation, and other important variables. Following Alig (1986) and Miller and Plantinga (1999), land rent provides guiding principle for selection of independent variables in a regional model of land use acreages. Because of insufficient series of land rent data, I employ county-level averages, which may include relevant proxy variables for the land quality characteristics of the county. The proxies include real net income from land products or other measures that influence land rent (e.g., land value, population density) and which would cause significant shifts in land rent profiles. Aggregate land-use shares depend on the land quality allocation, net returns to alternative land uses, and coefficients measure the effect of explanatory variables of expected shares of the county (Plantinga 1996, Chakir and Le Gallo, 2013). To examine the associations between land use drivers and land use dynamics, multinomial logistic share equation can be used here.

$$P_{kit} = \frac{\exp(\beta_k X_{kit})}{\sum_{i=1}^{k} \exp(\beta_i X_{kit})} \text{ for } k = 1,2, \ldots..K \tag{1}$$

This land use share model equation represent the land share P allocated to land use k, to total land in county i for i=1…..N, at time t for t = 1……T normalized with respect to a common land share category. P_{kit} is a non-stochastic vector in the equation that is county specific (land quality, land value) and land use specific (farm revenue and expenditure) variables that influence land use. When county level data are applied, land use shares are observed as percentage of total county area for the given uses and this model can be expressed as a linear and logarithmic equation normalized to a common share of the land use K.

$$y_{kit} = ln\left(\frac{P_{kit}}{P_{Kit}}\right) = (\beta_k - \beta_K)X_{it} + u_{kit} \tag{2}$$

y_{kit} is the expected land use share for land use k, in county i, at time t signifies the optimal land allocation and u_{kit} is the random error term with mean zero. Model (2) consists of K-1 equations that estimate the marginal effects of predictor variables on the log of normalized land use share (Bell et al., 2006).

Data and Methodology

The data are collected from the USDA NASS Census of Agriculture database which are the most recent 6 years dataset in 5-year intervals (1992, 1997, 2002, 2007, 2012 and 2017) for Alabama. County-level observations for Cropland, Pastureland, Urban land, Cotton, Corn, Peanut, and Soybean land shares are gathered from the Census years. Forest land and forest revenue data obtained for 1990, 2000, 2007, 2010, 2012, and 2016 from Forest Inventory and Analysis (FIA) database and since forest area tends to fluctuate less than agricultural land area, this data series interpolated to produce a consistent set of observations.

Four major land use shares (Crop, Pasture, Urban and Forest), and five crop shares (Cotton, Corn, Soybean, Peanut and other crops) were considered as dependent variables in this study. For the predictor variables, farm revenue, forest revenue, agricultural land value, and government payments are calculated per acre and population density (total population divided by total land area in the county) and Irrigated land share are obtained for all census years. Total population for each county is taken by the decadal Census of Population and Housing conducted by Bureau of the Census, and linear interpolation is used to produce estimates corresponding to the Census of Agriculture years. Two land quality variables are included in the model to control for differences across counties from Soil Survey Geographical database (SSURGO). LCC ratings are derived from the county-level USDA web soil surveys and based on 12 soil characteristics (e.g., slope and permeability). Ratings range from I to VIII, where I is the most productive land and VIII is the least productive. The land quality variables are not indexed by time and the measures will remain constant over time. Also, from the U.S. Geological Survey (USGS) county-level data (1990, 1995, 2000, 2005, 2010 and 2015), five variables were obtained. Irrigation groundwater withdrawals (fresh), Irrigation, surface-water withdrawals (fresh), Irrigated acres (sprinkler), Irrigated acres (micro-irrigation), and Irrigated acres (surface - flood). These data set also interpolated into census years to produce a consistent set of observations. In addition to the independent variables described above, 1992 to 2017 survey years used to capture the time trend and spatial heterogeneity across counties. Also, district fixed effects has applied in the estimation and these agricultural districts are classified into six regions, which are blackbelt, coastal plains, mountain, northern, upper plains and wiregrass. Table 1.1 reports the summary statistics of the variables used in the empirical analysis.

Table 1.1. Summary statistics for all the variables included in the model

Variable	Obs	Mean	Std. Dev.	Min	Max
Crop land share	402	0.1141	0.0851	0.0098	0.5059
Pasture land share	402	0.0552	0.0408	0.0044	0.2323
Urban land share	402	0.1381	0.0828	0.0015	0.4280
Forest land share	402	0.6927	0.1501	0.2336	0.9501
Cotton land share	402	0.0142	0.0237	0	0.1757
Corn land share	402	0.0081	0.0123	0	0.0777
Soybean land share	402	0.0087	0.0181	0	0.1373
Peanut land share	402	0.0149	0.0247	0	0.1137
Other crops share	402	0.9540	0.0483	0.7641	1
Farm Revenue ($/acre)	402	0.1397	0.1735	0	1.1527
Farm Expenditure ($/acre)	402	0.1134	0.1322	0.0022	0.8637
Forest Revenue ($/acre)	402	18.6550	19.3381	0	218.4543
Agricultural Land Value ($/acre)	402	0.5766	0.4211	0.5809	2.8006
Population Density	402	0.1342	0.1557	0.0188	0.9339
Government Payments ($/acre)	402	2.7145	3.3252	0.0360	22.1135
Average land quality	402	3.7674	0.4494	2.9231	4.9091
Proportion of land in high quality	402	0.2133	0.1157	0.026	0.53

Irrigated Land share	402	0.0034	0.0055	0	0.0366
Irrigated fresh ground water withdrawals	402	0.8851	3.7983	0	47.17
Irrigated fresh surface water withdrawals	402	1.1513	1.7341	0	11.11
Irrigated sprinkler acreage	402	1838.134	3206.84	0	35710
Irrigated micro irrigation acreage	402	19.403	74.1512	0	940
Irrigated surface (flood) acreage	402	3.1094	20.8673	0	230

In this study, I investigate the effects of land, socioeconomic determinants and irrigation water use on land use shares. In a separate analysis, I also examine the impacts of land, socioeconomic and irrigation factors on different crop shares in Alabama. The basic regression model used with relevant vectors is follows.

$$y_{it} = \beta_0 + \beta X_{it} + T_t + \theta_t + \varepsilon_{it} \tag{3}$$

where, y_{it} is land use share / crop share in county i and year t, β_0 is the constant, X_{it} represent the set of independent variables, T_t contains the time trend, θ_t contains district fixed effect and ε_{it} is the error term. Further, this model equation can be specified as follows.

$$L_{it} = \beta_0 + \alpha L_{i,t-1} + \beta X_{it} + T_t + \theta_t + \varepsilon_{it} \tag{4}$$

$$C_{it} = \beta_0 + \alpha C_{i,t-1} + \beta X_{it} + T_t + \theta_t + \varepsilon_{it} \tag{5}$$

L_{it} is the land use share (dependent variable) considering four land uses; crop, pasture, urban and forestry allocated by county i in time t, $L_{i,t-1}$ is the lagged dependent variable, X_{it} is the set of explanatory variables (farm revenue, farm expenditure, forest revenue, land value, population density, government payments, land quality factors, and irrigation factors). C_{it} is the crop share of one of the five types of crops: cotton, corn, soybean, peanut and other crops.

$C_{i,t-1}$ is the lagged dependent variable of crop share and explanatory variables, time trend T_t and district fixed effects θ_t and error term same as land use model equations.

Ordinary least squares (OLS) method is the best estimation method to apply to a dynamic panel regression models like equations (5) and (6), if the model satisfies all the classical assumptions of OLS. However, Nickel (1981) explained that these dynamic panel models can be biased due to the correlation between the lagged dependent variable and county fixed effects. This association violates the strict exogeneity assumption as the OLS estimator is biased and inconsistent. Transforming data and removing fixed effects will solve this problem but within group transformation, the lagged dependent variable remains correlated and predetermined explanatory variables (not strictly exogenous) can be correlated with the error term (Roodman, 2009a). Hence, the fixed-effects (FE) estimator is biased and inconsistent with this condition, estimated coefficient of the lagged dependent variable is downward biased in the case of the FE estimator and positive and upward biased on OLS estimator (Roodman 2009a, 2009b). Arellano and Bond (1991) established a more efficient estimator called difference generalized method of moments (GMM), utilizing all properly lagged endogenous and other exogenous variables as instruments to estimate a dynamic panel model (Roodman 2009a). Arellano-Bond estimator controls for the endogeneity of the lagged dependent variable and removes any fixed effects that show up by taking first difference. Blundell and Bond (1998) also formed an estimator named system GMM to overcome overidentification and dynamic panel bias problems.

When land use and crop share equations as specified in equations (5) and (6), all the equations for land use and crop shares are independent of each other and can be estimated separately and the error terms of the equations contemporaneously correlated with each other. Although these models can be estimated equation-by equation using OLS, they are not efficient as seemingly unrelated regression (SUR) model which use specific form of the variance-covariance matrix with feasible generalized least squares. The SUR model was introduced by Zellner (1962) to estimate the parameters of the system, accounting for heteroskedasticity and correlation in the contemporaneous errors through equations.

In this paper, I consider OLS and FE land use and crop share models for dynamic panel with cluster-robust standard errors using data for counties in all available years as a base. To address the problems for correlation and fixed effects, then I estimate land use and crop share functions using SUR model. Finally, I estimate the model with the Arellano-Bond two-step difference GMM dynamic panel data model where lagged-dependent variables are included on the right-hand side of the model for all land use and crop share equations. In addition, to eliminate the multicollinearity among explanatory variables, principal component analysis (PCA) developed by Pearson (1901) is considered in land use change and cropping patterns analysis in this paper. The PCA reduces the dimensionality by converting a set of correlated features into a set of values of linearly uncorrelated variables using orthogonal transformation without losing much information. The PCA method allows to remove the inter -dependence (collinearity) of the explanatory variables and create new variables called principal components (PC) as linear combinations of the original variables.

Results

Results of the empirical analysis for land use shares are shown in Table 1.2 to Table 1.5. Table 1.2 and 1.3 show the results for major land use shares using OLS and SUR estimation approaches and examine the effect of each of the land use determinants and irrigation variables on land use shares. These effects are interpreted as the changes in the land use shares for one-unit changes in the independent variables, all else equal. In dynamic panel OLS and SUR regression models, lagged share coefficients show significant and positive effects on all land use shares. The land determinants results reveal that farm revenue and farm expenses have the expected effects on all land use shares. Farm revenue shows positive effects on crop and forest land shares and negative effects on pasture and urban land shares. Farm expenses have positive significant effects on pastureland and urban land shares, negative effects on crop and forestland shares. Forest revenue show positive significant effect on pastureland and negative effects in crop and urban land shares and not significant with forest land. Agriculture land value also show positive significant effects on crop and pasture lands and negative effect on forest land. Population density show positive effect on urban land and negative effect on forest land shares and Government payments play significant positive role on crop land share and negative role on urban lands. Average land quality is not showing any significant effect with lands but proportion of high quality land show positive effects on crop and pasture land shares. These models shows that an increase in irrigation share increases cropland use but has a negative significant effect on urban land shares. Fresh ground water withdrawals show negative effect on urban lands and positive effect on forest land. Fresh surface water withdrawals also has significant positive effect on crop land shares.

As irrigation methods, sprinkler and micro irrigation using land show significant effects on pasture, urban and forest land shares. The time trend indicates negative and significant effect on crop land share and pasture, forest land shares show positive and significant effects. There are significant district fixed effects on crop, urban and forest land shares in the OLS regression model.

Table 1.2. Estimation Results for Land use shares of dynamic OLS response model

Variable	OLS			
	Crop	Pasture	Urban	Forest
Lagged Dependent variable	0.7660***	0.8092***	0.2731***	0.5262***
	(0.0309)	(0.0296)	(0.0529)	(0.0451)
Farm Revenue	0.1772***	-0.1202**	-0.4701***	0.4643***
	(0.0519)	(0.0390)	(0.1414)	(0.1585)
Farm Expenses	-0.2549***	0.1600**	0.7398***	-0.6746***
	(0.0683)	(0.0513)	(0.1857)	(0.2083)
Forest Revenue	-0.0002*	0.0003***	-0.0005**	0.0004
	(0.0001)	(0.00007)	(0.0002)	(0.0003)
Agricultural land value	0.0224***	0.0256***	-0.0114	-0.0799***

	(0.0066)	(0.0052)	(0.0164)	(0.0202)
Population Density	-0.0085	0.0047	0.1992***	-0.1046***
	(0.0082)	(0.0062)	(0.0273)	(0.0273)
Government payments	0.0025***	0.0006	-0.0035**	-0.0007
	(0.0006)	(0.0004)	(0.0015)	(0.0016)
Average land quality	0.0028	-0.0024	-0.0151	0.0113
	(0.0034)	(0.0025)	(0.0092)	(0.0103)
Proportion of land in high quality	0.0361**	0.0311***	0.0119	-0.0467
	(0.0162)	(0.0116)	(0.0419)	(0.0475)
Irrigated Land share	0.3924***	0.1336	-1.9857**	0.9006
	(0.3744)	(0.2698)	(0.9697)	(1.0945)
Fresh ground water withdrawals	0.0007	0.0005	-0.0040***	0.0035**
	(0.0004)	(0.0003)	(0.0013)	(0.0014)
Fresh surface water withdrawals	0.0026***	-0.0010	0.0025	-0.0039
	(0.0009)	(0.0006)	(0.0023)	(0.0027)
Sprinkler acreage	-0.0013	-0.0023***	0.0069**	-0.0029
	(0.0011)	(0.0008)	(0.0031)	(0.0035)
Micro irrigation acreage	-0.0233	0.0235*	0.0809*	-0.0980*
	(0.0177)	(0.0133)	(0.0482)	(0.0539)
Surface (flood) acreage	-0.0973	-0.0070	0.0401	0.0832
	(0.0629)	(0.0472)	(0.1707)	(0.1915)
Time Trend	-0.0049***	0.0026***	0.0029	0.0088***
	(0.0013)	(0.0008)	(0.0028)	(0.0034)
Agricultural District Fixed effects	Yes	No	Yes	Yes
Constant	0.0109	0.0254	0.1252***	0.3249***
	(0.0169)	(0.0119)	(0.0449)	(0.0556)
R^2	0.9498	0.8859	0.6253	0.8547

Notes: ***, ** and * indicate significance levels 1%, 5% and 10% respectively. Standard errors for coefficients in parentheses.

Table 1.3. Estimation Results for Land use shares of dynamic SUR response model

Variable	SUR			
	Crop	Pasture	Urban	Forest
Lagged Dependent variable	0.7167***	0.7525***	0.6145***	0.6651***
	(0.0261)	(0.0250)	(0.0288)	(0.0260)
Farm Revenue	0.1759***	-0.1217***	-0.5216***	0.4784***
	(0.0519)	(0.0390)	(0.1412)	(0.1584)
Farm Expenses	-0.2527***	0.1598***	0.7668***	-0.6799***
	(0.0683)	(0.0513)	(0.1856)	(0.2083)
Forest Revenue	-0.0001*	0.0002***	-0.0005**	0.0004
	(0.0001)	(0.0007)	(0.0002)	(0.0003)
Agricultural land value	0.0268***	0.0304***	-0.0146	-0.0544***
	(0.0065)	(0.0050)	(0.0164)	(0.0190)
Population Density	-0.0097	0.0029	0.0971***	-0.0711***
	(0.0082)	(0.0062)	(0.0239)	(0.0258)

Government payments	0.0027***	0.0005	-0.0035**	-0.0004
	(0.0005)	(0.0004)	(0.0014)	(0.0016)
Average land quality	0.0023	-0.0025	-0.0064	0.0061
	(0.0033)	(0.0025)	(0.0091)	(0.0103)
Proportion of land in high quality	0.0441***	-0.0324***	0.0081	-0.0257
	(0.0160)	(0.0116)	(0.0419)	(0.0472)
Irrigated Land share	0.5726***	0.0730	-2.0066**	1.2685
	(0.3695)	(0.2693)	(0.9697)	(1.0901)
Fresh ground water withdrawals	0.0007	0.0005	-0.0050***	0.0039***
	(0.0005)	(0.0003)	(0.0013)	(0.0014)
Fresh surface water withdrawals	0.0027***	-0.0009	0.0005	-0.0023
	(0.0008)	(0.0006)	(0.0023)	(0.0026)
Sprinkler acreage	-0.0015	-0.0025***	0.0088***	-0.0046
	(0.0011)	(0.0008)	(0.0031)	(0.0034)
Micro irrigation acreage	-0.0249	0.0249*	0.1048**	-0.1088**
	(0.0177)	(0.0133)	(0.0481)	(0.0539)
Surface (flood) acreage	-0.1004	-0.0058	0.0269	0.0827
	(0.0629)	(0.0472)	(0.1707)	(0.1915)
Time Trend	-0.0063***	-0.0023***	0.0047*	0.0048
	(0.0013)	(0.0008)	(0.0028)	(0.0032)
Agricultural District Fixed effects	Yes	No	Yes	No
Constant	0.0197	0.0255**	0.0496	0.2418***
	(0.0166)	(0.0119)	(0.0438)	(0.0511)
R^2	0.9494	0.8845	0.5761	0.8504

Notes: ***, ** and * indicate significance levels 1%, 5% and 10% respectively. Standard errors for coefficients in parentheses.

Tables 1.4 and 1.5 present the results for Alabama by major land use shares for fixed effects model and Arellano-bond (AB) estimator approach and show the effects of each of the land use determinants and irrigation variables on land use shares. These effects are interpreted as the changes in the land use shares for one-unit changes in the independent variables, all else equal. Table 1.4 presents the results of land use shares dynamic panel fixed effects model.

Table 1.4. Estimation Results for Land use shares of dynamic fixed effect model

Variable	Fixed Effects			
	Crop	Pasture	Urban	Forest
Lagged Dependent variable	0.2015**	0.2877***	-0.0237	0.0293
	(0.0895)	(0.0748)	(0.0698)	(0.0679)
Farm Revenue	0.1577***	-0.1291**	-0.3994***	0.3923**
	(0.0545)	(0.0587)	(0.1375)	(0.1571)
Farm Expenses	-0.2907***	0.2159**	0.7651***	-0.7340***
	(0.0722)	(0.0917)	(0.1847)	(0.2139)
Forest Revenue	-0.0002	0.0013**	-0.0005**	0.0005**
	(0.0001)	(0.00005)	(0.0002)	(0.0002)
Agricultural land value	0.0029	0.0299**	-0.1321***	0.0908***
	(0.0148)	(0.0125)	(0.0296)	(0.0302)
Population Density	-0.0740	0.0218	1.0251***	-0.9031***
	(0.0668)	(0.0402)	(0.2583)	(0.2309)
Government payments	-0.0003	0.0008*	-0.0039**	0.0032**
	(0.0005)	(0.0005)	(0.0016)	(0.0016)
Average land quality	0.1372***	0.0288***	-0.0865**	-0.2462***
	(0.0137)	(0.0064)	(0.0415)	(0.0571)
Irrigated Land share	-0.1421	0.0437	-1.6862*	1.7973
	(0.7647)	(0.5391)	(0.9642)	(1.1504)
Fresh ground water withdrawals	0.0007	0.0002	-0.0036***	0.0027**
	(0.0005)	(0.0002)	(0.0012)	(0.0012)
Fresh surface water withdrawals	0.0025**	-0.0017**	0.0005	-0.0014
	(0.0011)	(0.0007)	(0.0047)	(0.0047)
Sprinkler acreage	-0.0011	-0.0017	0.0075**	-0.0048
	(0.0018)	(0.0013)	(0.0034)	(0.0034)
Micro irrigation acreage	-0.0046	-0.0008	0.1009	-0.0953
	(0.0192)	(0.0103)	(0.0613)	(0.0597)
Surface (flood) acreage	-0.1111	-0.0064	-0.0021	0.1267
	(0.0954)	(0.0371)	(0.0849)	(0.1089)
Time Trend	-0.0074***	0.0019*	0.0106**	-0.0051
	(0.0018)	(0.0010)	(0.0044)	(0.0044)
Agricultural District Fixed effects	Yes	Yes	Yes	Yes
Observations	335	335	335	335
R^2	0.6700	0.6966	0.4695	0.3824

Notes: ***, ** and * indicate significance levels 1%, 5% and 10% respectively. Standard errors for coefficients in parentheses.

Table 1.5 shows the results for AB estimator and shows the land use share using two-step difference GMM model and to control for within correlation standard errors are grouped as clusters by counties.

Table 1.5. Estimation Results for Land use shares of dynamic Arellano-Bond model

Variable	Arellano-Bond estimator			
	Crop	Pasture	Urban	Forest
Lagged Dependent variable	0.2746	0.3084	-0.0126	0.1010
	(0.1556)	(0.1293)	(0.1147)	(0.1548)
Farm Revenue	0.1561*	-0.1397**	-0.1797	0.2564
	(0.0886)	(0.0645)	(0.2161)	(0.3041)
Farm Expenses	-0.3897***	0.2399**	0.6661***	-1.1544**
	(0.0816)	(0.1092)	(0.1834)	(0.5570)
Forest Revenue	-0.0004**	0.0001	0.0002	-0.0001
	(0.0001)	(0.0007)	(0.0004)	(0.0007)
Agricultural land value	0.0115	0.0393	-0.2792***	0.3627**
	(0.0387)	(0.0259)	(0.0997)	(0.2783)
Population Density	-0.7768*	0.2458	4.7482**	-4.9875
	(0.4162)	(0.3213)	(1.8842)	(3.0855)
Government payments	0.0002***	0.0014	-0.0042**	-0.0064
	(0.0014)	(0.0010)	(0.0058)	(0.0059)
Average land quality	0.2085***	0.0267**	0.0038	-0.3667
	(0.0397)	(0.0774)	(0.2443)	(0.3961)
Irrigated Land share	0.6467**	-0.2214	-1.9918**	2.1423
	(1.0934)	(0.7401)	(2.7140)	(2.4711)
Fresh ground water withdrawals	0.0002**	0.0005	-0.0037*	0.0008
	(0.0009)	(0.0007)	(0.0024)	(0.0026)
Fresh surface water withdrawals	0.0075**	-0.0021	0.0016	0.0078
	(0.0028)	(0.0016)	(0.0064)	(0.0074)
Sprinkler acreage	-0.0004	-0.0033*	-0.0078	0.0112
	(0.0024)	(0.0018)	(0.0083)	(0.0154)
Micro irrigation acreage	-0.0012	0.0059	-0.0053	-0.0076
	(0.0583)	(0.0226)	(0.0545)	(0.0754)
Surface (flood) acreage	-0.0069	-0.0247	-0.0385	0.0599
	(0.1813)	(0.1163)	(0.5709)	(0.3295)
Time Trend	-0.0055*	0.0004	0.0102	-0.0085
	(0.0031)	(0.0014)	(0.0075)	(0.0132)
Agricultural district fixed effects	Yes	Yes	Yes	Yes
Observations	268	268	268	268
R^2	-	-	-	-
Arellano-Bond test for AR(1): p-value	0.009	0.047	0.049	0.045
Arellano-Bond test for AR(1): p-value	0.310	0.395	0.298	0.983
Hansen test for overidentification: p-value	0.257	0.180	0.408	0.818

Notes: ***, ** and * indicate significance levels 1%, 5% and 10% respectively. Standard errors for coefficients in parentheses

The coefficients associated with the lagged land use shares are not significantly different from zero in all four land use share equations in AB model. But show positive significant effects for Crop and Pastureland shares in the fixed effects model. Farm revenue shows positive significant effect on crop land and negative effect on pasture land shares. The farm revenue coefficient of 0.1561, which is significant at the 10% significance level indicates that a 10% rise in farm revenue results in an increase of crop land share by 1.5%, all else. Also, 10% increase in farm revenue is associated with a decrease of pastureland share by 1.4% compared to other land shares. Farm expenses also play a significant role in land use shares with positive effects of pasture and urban land shares and negative effects on crop and forest land shares. Forest revenue is also important as an increase in forest revenue is associated with a decrease in crop land shares which means that forest industry has significantly impacted farmland in last two decades. Agricultural land values have negative effects on urban lands and positive effects on forest lands. Population density is associated with a higher share of urban land use and lower share of crop land. Government payments are positively associated with crop land shares and negatively with urban land shares. Average land quality is positive and significant for crop and pasture land shares. Irrigation share is positive for cropland use and negative for urban land use. Fresh ground water withdrawals show negative effect on urban land share and positive effect on crop lands. Fresh surface water withdrawals also have significant positive effect on crop land shares. As for different irrigation methods, sprinkler irrigation show significant negative effects on pastureland shares. The time trend implies that crop land shares have been decreasing for the last 25 years. Also, all land use shares have significant Alabama agricultural district fixed effects. I compared difference GMM estimates with the results obtained from other estimators (FE, SUR and OLS) and conducted sensitivity analysis and these models confirm that AB estimates are unbiased, consistent, and robust with the dynamic panel model.

Tables 1.6 and 1.7 show the OLS and SUR models and show the effect of each of the land use determinants and irrigation variables on crop shares.

Table 1.6. Estimation Results for Crop shares of OLS response model

Variable	OLS				
	Cotton	Corn	Soybean	Peanut	Other
Lagged Dependent variable	0.1516***	0.6852***	0.2241***	0.5678***	0.4258***
	(0.0550)	(0.0371)	(0.0621)	(0.0420)	(0.0501)
Farm Revenue	-0.0326	-0.0509***	0.0314**	-0.0159	0.0684
	(0.0426)	(0.0148)	(0.0499)	(0.0514)	(0.0914)
Farm Expenses	0.0487	0.0541***	-0.0574	0.0231	-0.0781
	(0.0560)	(0.0194)	(0.0657)	(0.0675)	(0.1202)
Forest Revenue	-0.0007	0.0003	0.0003	-0.0001*	0.0002
	(0.0006)	(0.0002)	(0.0009)	(0.0009)	(0.0001)
Agricultural land value	-0.0114**	0.0038**	0.0103*	-0.0038	-0.0003
	(0.0049)	(0.0018)	(0.0058)	(0.0060)	(0.0106)
Population Density	0.0041	0.0006	-0.0086	-0.0090	0.0148
	(0.0067)	(0.0024)	(0.0079)	(0.0081)	(0.0144)
Government payments	0.0021***	0.0011***	-0.0008	0.0002	-0.0023**
	(0.0005)	(0.0001)	(0.0005)	(0.0005)	(0.0009)
Average land quality	-0.0035	0.0004	-0.0006	-0.0019	0.0059
	(0.0027)	(0.0009)	(0.0032)	(0.0033)	(0.0059)
Proportion of land in high quality	0.0316**	-0.0016	-0.0029	0.0102	-0.0366
	(0.0127)	(0.0045)	(0.0148)	(0.0153)	(0.0273)
Irrigated Land share	1.6220***	-0.0894	0.0263	-0.2655	-1.1693*
	(0.2956)	(0.1026)	(0.3437)	(0.3536)	(0.6352)
Fresh ground water withdrawals	0.0004	0.0002*	-0.0003	0.0005	-0.0009
	(0.0003)	(0.0002)	(0.0004)	(0.0005)	(0.0008)
Fresh surface water withdrawals	-0.0003	0.0003	-0.0002	-0.0005	0.0013
	(0.0007)	(0.0002)	(0.0008)	(0.0008)	(0.0015)
Sprinkler acreage	-0.0003**	-0.0002	0.0012	0.0003	-0.0013
	(0.0009)	(0.0003)	(0.0011)	(0.0011)	(0.0019)
Micro irrigation acreage	-0.0289**	-0.0085*	-0.0157	-0.0096	0.0662**
	(0.0145)	(0.0051)	(0.0170)	(0.0175)	(0.0314)
Surface (flood) acreage	-0.0744	-0.0487***	-0.0507	-0.0258	0.1991*
	(0.0516)	(0.0182)	(0.0605)	(0.0621)	(0.1110)
Time Trend	-0.0008	-0.0006**	-0.0005	0.0007	0.0008
	(0.0009)	(0.0003)	(0.0009)	(0.0010)	(0.0018)
Agricultural district effects	Yes	Yes	No	No	Yes
Constant	0.0251*	0.0038	0.0100	0.0139	0.5238***
	(0.0131)	(0.0046)	(0.0153)	(0.0159)	(0.0537)
R^2	0.5633	0.8202	0.0961	0.3641	0.5352

Notes: ***, ** and * indicate significance levels 1%, 5% and 10% respectively. Standard errors for coefficients in parentheses.

Table 1.7. Estimation Results for Crop shares of SUR response model

Variable	SUR				
	Cotton	Corn	Soybean	Peanut	Other
Lagged Dependent variable	0.3970***	0.6116***	0.4332***	0.4640***	0.4559***
	(0.0295)	(0.0325)	(0.0331)	(0.0279)	(0.0264)
Farm Revenue	-0.0313	-0.0518***	0.0285***	-0.0138	0.0698
	(0.0426)	(0.0148)	(0.0499)	(0.0513)	(0.0913)
Farm Expenses	0.0539	0.0551***	-0.0538	0.0214	-0.0800
	(0.0560)	(0.0195)	(0.0657)	(0.0675)	(0.1202)
Forest Revenue	-0.0009	0.0003	0.0003	-0.0002**	0.0002
	(0.0007)	(0.0002)	(0.0008)	(0.0009)	(0.0001)
Agricultural land value	-0.0121**	0.0046***	0.0108**	-0.0045	-0.0005
	(0.0049)	(0.0017)	(0.0058)	(0.0059)	(0.0106)
Population Density	0.0029	-0.0002	-0.0064	-0.0090	0.0143
	(0.0067)	(0.0023)	(0.0079)	(0.0081)	(0.0145)
Government payments	0.0011**	0.0011***	-0.0002	0.0003	-0.0021**
	(0.0004)	(0.0001)	(0.0005)	(0.0005)	(0.0009)
Average land quality	-0.0033	0.0004	-0.0007	-0.0021	0.0057
	(0.0027)	(0.0009)	(0.0032)	(0.0033)	(0.0059)
Proportion of land in high quality	0.0248**	0.0001	-0.0035	0.0109	-0.0345
	(0.0127)	(0.0045)	(0.0148)	(0.0153)	(0.0272)
Irrigated Land share	1.4352***	-0.0634	0.0290	-0.2409	-1.1108**
	(0.2933)	(0.1025)	(0.3434)	(0.3531)	(0.6298)
Fresh ground water withdrawals	0.0005	0.0003**	-0.0004	0.0005	-0.0009
	(0.0003)	(0.0001)	(0.0004)	(0.0005)	(0.0008)
Fresh surface water withdrawals	-0.0008	0.0002	0.0002	-0.0006	0.0013
	(0.0007)	(0.0002)	(0.0008)	(0.0008)	(0.0015)
Sprinkler acreage	-0.0004	-0.0001	0.0013	0.0002	-0.0013
	(0.0009)	(0.0003)	(0.0011)	(0.0011)	(0.0019)
Micro irrigation acreage	-0.0247**	-0.0105**	-0.0132	-0.0115*	0.0633**
	(0.0143)	(0.0051)	(0.0170)	(0.0172)	(0.0311)
Surface (flood) acreage	-0.0587	-0.0555***	0.0422	0.0263	0.1927*
	(0.0515)	(0.0182)	(0.0604)	(0.0621)	(0.1106)
Time Trend	0.0003	-0.0007**	-0.0005	0.0006	0.0006
	(0.0008)	(0.0003)	(0.0009)	(0.0010)	(0.0018)
Agricultural district effects	Yes	Yes	No	No	Yes
Constant	0.0200	0.0042	0.0075	0.0161	0.4963***
	(0.0131)	(0.0045)	(0.0153)	(0.0158)	(0.0369)
R^2	0.5358	0.8180	0.0638	0.3619	0.5347

Notes: ***, ** and * indicate significance levels 1%, 5% and 10% respectively. Standard errors for coefficients in parentheses.

Tables 1.8 and 1.9 present the results for crop shares in Alabama for fixed effects model and Arellano-bond (AB) estimator models. Table 1.9 shows the results for AB estimator and shows the crop share using two-step difference GMM model and to control for within correlation standard errors are grouped as clusters by counties.

Table 1.8. Estimation Results for Crop shares of Fixed effects response model

Variable	Fixed Effects				
	Cotton	Corn	Soybean	Peanut	Other
Lagged Dependent variable	-0.3213**	0.4786***	0.0016	0.2759***	-0.0163
	(0.1301)	(0.0658)	(0.0983)	(0.0414)	(0.0565)
Farm Revenue	-0.0115	-0.0521***	0.0224	0.0067	0.0094
	(0.0423)	(0.0172)	(0.0622)	(0.0546)	(0.0897)
Farm Expenses	0.0321	0.0640***	-0.0581	-0.0064	-0.0015
	(0.0530)	(0.0231)	(0.0701)	(0.0799)	(0.1252)
Forest Revenue	-0.00002	9.12e-06	0.00002	0.0002**	0.0002
	(0.00004)	(0.00001)	(0.0001)	(0.0001)	(0.0002)
Agricultural land value	-0.0366**	0.0047	0.0284	-0.0034	0.0017
	(0.0178)	(0.0039)	(0.0215)	(0.0095)	(0.0324)
Population Density	0.0819	0.0157	-0.0126	0.0232	-0.1376
	(0.0754)	(0.0154)	(0.0652)	(0.0390)	(0.0947)
Government payments	0.0016	0.0012**	-0.0011**	-0.0011	0.0016
	(0.0012)	(0.0004)	(0.0005)	(0.0007)	(0.0015)
Average land quality	-0.0010	-0.0064	-0.0104	0.0042	0.0101
	(0.0073)	(0.0042)	(0.0089)	(0.0078)	(0.0143)
Irrigated Land share	0.8534*	0.2959**	0.5522	-0.4253	-1.2895
	(0.4981)	(0.1535)	(0.7986)	(0.6784)	(0.8886)
Fresh ground water withdrawals	0.0007***	0.0002	-0.0003	0.0012*	-0.0022**
	(0.0003)	(0.0002)	(0.0005)	(0.0006)	(0.0010)
Fresh surface water withdrawals	-0.0016*	0.0004	0.0009	-0.0024	0.0038
	(0.0008)	(0.0003)	(0.0011)	(0.0015)	(0.0025)
Sprinkler acreage	-0.0009	-0.0006	0.0014	0.0011	-0.0025
	(0.0008)	(0.0004)	(0.0023)	(0.0013)	(0.0027)
Micro irrigation acreage	0.0049	-0.0069	-0.0018	0.0099	0.0064
	(0.0142)	(0.0049)	(0.0104)	(0.0103)	(0.0287)
Surface (flood) acreage	0.0010	-0.0163	-0.0465	-0.0744	0.2564*
	(0.0333)	(0.0135)	(0.0533)	(0.0554)	(0.1531)
Time Trend	0.0019**	-0.0009**	-0.0021*	0.0007	0.0005
	(0.0009)	(0.0004)	(0.0012)	(0.0009)	(0.0022)
Agricultural district effects	Yes	Yes	Yes	Yes	Yes
Observations	335	335	335	335	335
R^2	0.2458	0.6451	0.0775	0.1745	0.0829

Notes: ***, ** and * indicate significance levels 1%, 5% and 10% respectively. Standard errors for coefficients in parentheses.

The coefficients associated with the lagged crop shares, corn, soybean, and peanut were significantly different from all five crop uses share equations. Farm revenue shows positive significant effect on soybeans and negative effects on cotton and corn crop shares.

Farm revenue increase of 10% is associated with an increase by 1% of soybean acreage and with 1.8% decrease and 0.4% decrease in cotton and corn acreage shares respectively. Farm expenses also show positive effects on corn share. Forest revenue is not associated with any of the crop shares. Agricultural land value shows negative effects on soybean lands and population density does not have a significant effect on crop shares. Government payments are positively associated with cotton and corn acreages. Proportion of high land quality is positive and significant for cotton and corn lands. Irrigation land share is positively associated with cotton, corn and soybean shares. Fresh ground water withdrawals are positive for cotton and peanut shares. As for irrigation methods, micro irrigation is negative for cotton and corn shares. Comparing difference GMM estimates with the results obtained from the FE, SUR and OLS models confirms that the estimated AB coefficients on the lagged dependent variables lie between the values obtained from OLS and FE estimators in tables 1.6 and 1.7 and dynamic panel estimates are valid in the model.

Table 1.9. Estimation Results for Crop shares of Arellano-Bond response model

Variable	Arellano-Bond estimator				
	Cotton	Corn	Soybean	Peanut	Other
Lagged Dependent variable	-0.1738	0.5388***	0.1016***	0.5491***	0.0976
	(0.1857)	(0.0936)	(0.0369)	(0.0688)	(0.1382)
Farm Revenue	-0.1882**	-0.0454***	0.1040**	-0.0466	-0.0098
	(0.0907)	(0.0137)	(0.0656)	(0.0558)	(0.1082)
Farm Expenses	0.0737	0.0555***	-0.0217	0.0844	-0.1343
	(0.0686)	(0.0137)	(0.0601)	(0.0971)	(0.1357)
Forest Revenue	0.0001	0.0001	6.34e-07	-0.0001	0.0001
	(0.0001)	(0.0002)	(0.00001)	(0.0001)	(0.0002)
Agricultural land value	0.0232	0.0038	-0.0371**	-0.0118	0.0498
	(0.0283)	(0.0063)	(0.0158)	(0.0153)	(0.0419)
Population Density	0.0140	-0.0014	-0.0041	-0.0057	0.0038
	(0.0110)	(0.0029)	(0.0090)	(0.0062)	(0.0169)
Government payments	0.0027**	0.0006**	-0.0006	0.0003	-0.0027
	(0.0013)	(0.0003)	(0.0009)	(0.0005)	(0.0019)
Average land quality	-0.0042	0.0002	-0.0008	-0.0037	0.0075
	(0.0053)	(0.0019)	(0.0028)	(0.0036)	(0.0101)
Proportion of land in high quality	0.0489**	0.0041*	0.0270	0.0126	-0.0841
	(0.0350)	(0.0075)	(0.0225)	(0.0180)	(0.0580)
Irrigated Land share	2.0192**	0.1155*	0.7715**	-0.1085	-2.8937**

	(0.9018)	(0.1962)	(0.5175)	(0.5754)	(1.1962)
Fresh ground water withdrawals	0.0003*	0.0002	-0.0003	0.0006*	-0.0016
	(0.0005)	(0.0001)	(0.0006)	(0.0005)	(0.0011)
Fresh surface water withdrawals	-0.0007	0.0003	0.0002	-0.0006	0.0025
	(0.0011)	(0.0002)	(0.0009)	(0.0005)	(0.0017)
Sprinkler acreage	0.0004	-0.0002	0.0017	-0.00002	-0.0030
	(0.0015)	(0.0004)	(0.0019)	(0.0011)	(0.0036)
Micro irrigation acreage	-0.0356**	-0.0066*	-0.0131	-0.0095	0.1134
	(0.0367)	(0.0056)	(0.0123)	(0.0089)	(0.0369)
Surface (flood) acreage	-0.0773	-0.0349	-0.0312	-0.0099	0.3380**
	(0.1247)	(0.0229)	(0.0528)	(0.0462)	(0.1522)
Time Trend	-0.0009	-0.0005	0.0008	0.0011	0.0016
	(0.0012)	(0.0003)	(0.0009)	(0.0010)	(0.0019)
Agricultural district effects	Yes	Yes	Yes	Yes	Yes
Observations	268	268	268	268	268
Arellano-Bond test for AR(1): p-value	0.008	0.007	0.027	0.000	0.006
Arellano-Bond test for AR(1): p-value	0.119	0.107	0.268	0.748	0.432
Hansen test for overidentification: p-value	0.221	0.381	0.803	0.163	0.179

Notes: ***, ** and * indicate significance levels 1%, 5% and 10% respectively. Standard errors for coefficients in parentheses

Robustness Checks

In what follows, I describe using Principal component analysis (PCA) to check for the sensitivity of results reported in Tables 1.10 and 1.11. Correlations between independent variables are shown in Figure 1.1, with high correlation coefficients close to 0.7 and above suggesting a multicollinearity problem between those variables. Also, Figure 1.1 illustrates the coefficient values through a heat plot representing low to high values as lighter colors to darker color squares. Farm expenses and Farm revenue show high correlation coefficients. When these explanatory variables are correlated with each other, there will be some drawbacks due to the fact that high correlations between predictor variables can have complications in the correct analysis. This problem can be removed through the PCA application. This is important to re-orient the data so that multitude of original variables summarized with fewer components that capture the maximum possible information and variation of original variables.

PCA creates new variables, called the principal components (PC), that are orthogonal and uncorrelated. These variables are linear combinations of the original variables. The PCA was performed to reduce the multicollinearity, and Table 1.10 presents all the principal components and their eigen values for all predictor variables in the model. According to the results of PCA (Table 1.10), there are four principal components out of fourteen with eigenvalues greater than 1 which were selected because eigenvalues represent variances and a component with an eigenvalue of less than 1 is not significant. Figure 1.2 also shows the scree plots of the eigenvalues of the PCA and according to the scree plot, PC1, PC2, PC3 and PC4 were selected as the regression variables. The first four principal components, explain 70.68% of the total variation, which should be sufficient for almost any application.

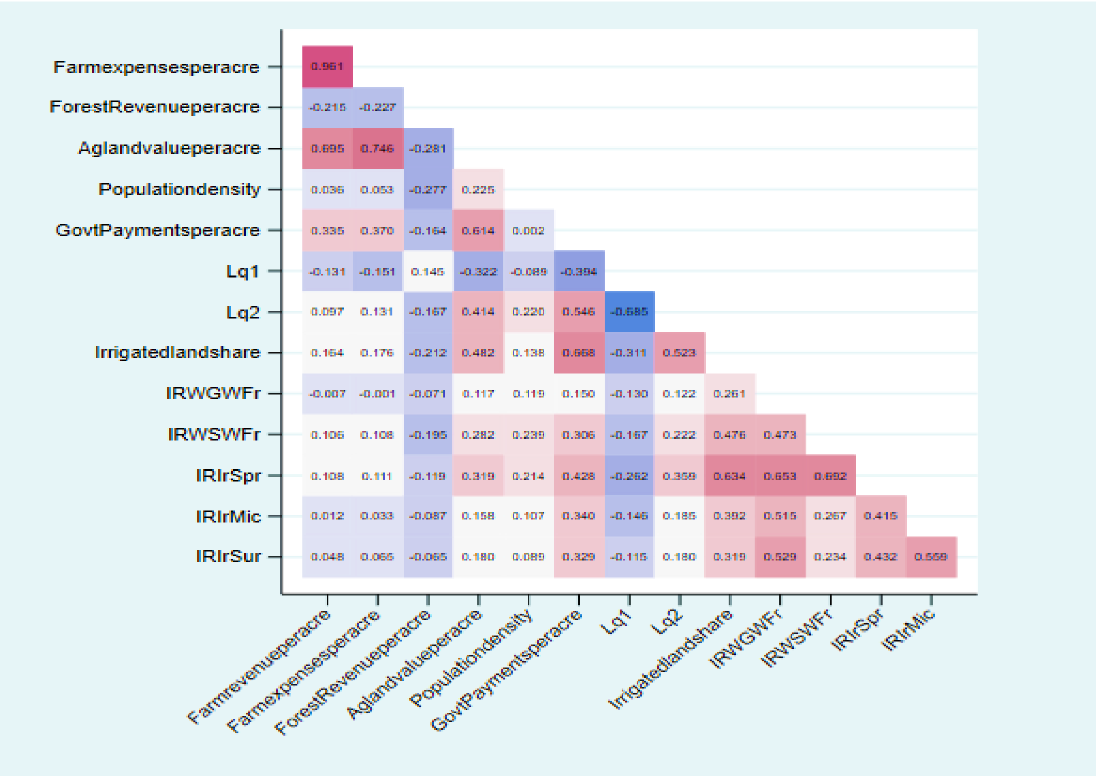

Figure 1.1. Heat plot of the correlated independent variables in the model

Table 1.10. All principal components and eigen values

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	4.7531	2.3268	0.3395	0.3395
Comp2	2.4263	0.9406	0.1733	0.5128
Comp3	1.4857	0.2555	0.1061	0.6189
Comp4	1.2301	0.3121	0.0879	0.7068
Comp5	0.9281	0.1566	0.0663	0.7731
Comp6	0.7715	0.0777	0.0551	0.8282
Comp7	0.6938	0.2485	0.0496	0.8778
Comp8	0.4453	0.1035	0.0318	0.9096
Comp9	0.3417	0.0527	0.0244	0.9340
Comp10	0.2890	0.0512	0.0206	0.9546
Comp11	0.2378	0.0296	0.0170	0.9716
Comp12	0.2083	0.0539	0.0149	0.9865
Comp13	0.1543	0.1193	0.0110	0.9975
Comp14	0.0350	.	0.0025	1.0000

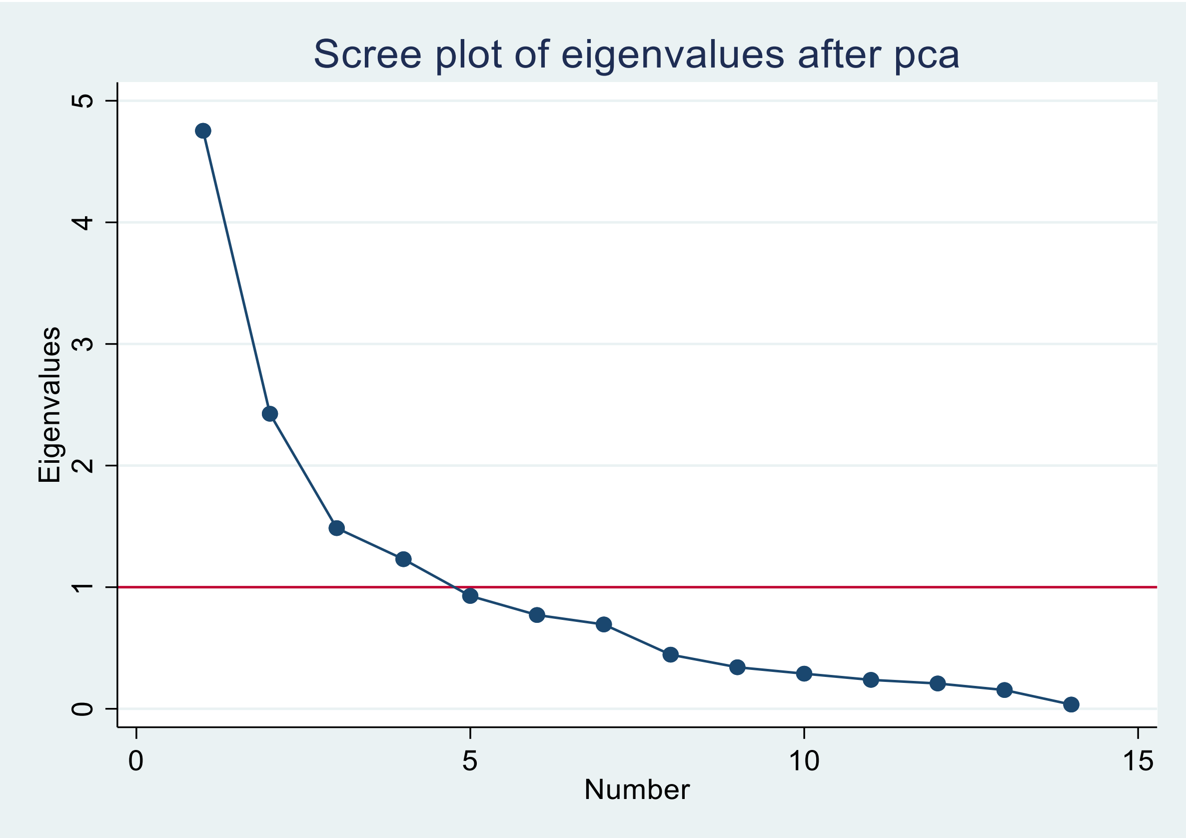

Figure 1.2 Scree plot of eigenvalues

The PCA indicated that the first four components account for the maximum estimated variation and they were extracted individually based on the eigenvalues (Table 1.11 below) revealing the pattern and principal components analysis of the data. Applying PCA, the variables were transformed into four principal components as shown in Table 1.11, maintaining all information of the original data. After the variable transformation, rotation was used to maximize the loadings of predictor variable on each principal component. The results show the rotation on the four PCs and the variance for each component and the values in bold correspond to the main contributions of the explanatory variables in each principal component. The first principal component (PC1) has important contributions to five irrigation variables, namely irrigated fresh ground water withdrawals, irrigated fresh surface water withdrawals, irrigated sprinkler acreage, irrigated micro irrigation acreage, and irrigated surface (flood) acreage with high positive loadings. PC2 is heavily loaded on Farm Revenue, Farm Expenditure, and Agricultural Land Value. PC3 also has important positive contributions to two land quality variables (Average Land quality and Proportion of high quality land), Government payments, and irrigated land share. The fourth principal component, PC4 is accounted for high negative loadings of forest revenue and positive for population density. These four components correspond to the eigen value variances of 3.1766, 2.7347, 2.5707 and 1.4133 (22.69%, 19.53%, 18.36%, and 10.10% of total variation present), accounting for 70.68% of total variation.

Table 1.11. Selected Principal components factor rotated loadings.

Model Variables	Component 1	2	3	4
	Eigen values			
	3.1766	2.7347	2.5707	1.4133
	Eigen vector loadings			
Farm Revenue	-0.0192	**0.6072**	-0.0793	-0.0151
Farm Expenses	-0.0195	**0.6093**	-0.0575	-0.0135
Forest Revenue	0.0440	-0.1151	-0.0175	**-0.5399**
Ag Land Value	0.0296	**0.4423**	0.1721	0.0807

Population Density	-0.0062	-0.0501	-0.0028	**0.6944**
Govt. Payments	0.1300	0.1844	**0.3884**	-0.1938
Avg. Land quality	0.0853	0.0622	**-0.5535**	0.0097
Proportion of high quality land	-0.0546	-0.0798	**0.6013**	0.0537
Irrigated land share	0.2161	0.0354	**0.3355**	0.0230
Ir. Fresh ground water withdrawals	**0.5051**	-0.0461	-0.1374	0.0491
Ir. Fresh surface water withdrawals	**0.3136**	-0.0004	0.0084	0.3251
Ir. sprinkler acreage	**0.4143**	-0.0176	0.0918	0.1447
Ir. micro irrigation acreage	**0.4396**	-0.0141	-0.0038	-0.1392
Ir. surface (flood) acreage	**0.4519**	0.0239	-0.0432	-0.1728
Percentage variation	22.69	19.53	18.36	10.10
Cumulative variation	22.69	42.22	60.59	70.68

Notes: Bold value indicates the highest eigen vector for the corresponding variable among the four principal components.

Conclusion

This paper integrates the dynamic panel regression models using Arellano-Bond (AB) estimator comparing other models such as fixed effects (FE), seemingly unrelated regression (SUR) and ordinary least squares (OLS) models to investigate the effect of irrigation, land quality and other economic and socioeconomic determinants on land use change and cropping patterns in the Southeast. This study compares the AB outcomes with the SUR, FE, and OLS outcomes in the land use shares and crop shares models to provide a clearer estimate on the impact of irrigation on land use and crop patterns using county level data on Alabama. Also, this paper utilizes the principal component analysis to examine the relative contributions of predictor variables used to construct land use and crop shares in this region. Agricultural revenue shows positive effects on crop and forest lands, but forest revenues and expenses show negative effects crop and forest land shares. Agricultural land value, population density, and Government payments are significantly associated with land use shares. The estimated effect of average land quality and the proportion of land quality show positive effects on crop and pasture land shares. Irrigated land share is positively associated with cropland and negatively with urban land, yet more detailed irrigation variables remain insignificant likely because irrigation has been very small.

Overall, crop shares are less associated with the explanatory variables possible because of the distribution of crop acreage within the cropland with different local climates and soils. Yet, cotton and corn are more associated with land attributes compared to peanut and soybean shares. Based on the PCA, the five irrigation variables (irrigated fresh ground water withdrawals, irrigated fresh surface water withdrawals, irrigated sprinkler acreage, irrigated micro irrigation acreage, and irrigated surface (flood) acreage) are positively correlated with first principal component (PC1). PC2 is positively correlated with proportion of high quality land, Government payments, and irrigated land share and negatively with average land quality. The third PC is also positively correlated with agricultural revenue, farm net return, and agricultural land value. The fourth principal component is negatively correlated with forest revenue and population density confirming the initial assumption that economic, social, environmental factors can influence land use and cropping decisions. As a summary of these results, agriculture and forest revenues and expenses have important effects on cropland and pastureland share trends and these results support the findings in the earlier studies that demographic factors such as land value, population density, government payments, and land quality are significant for land use decisions. Irrigated land acreage plays a major role influencing cropland, pastureland, and urban land use allocations and these results also support the hypothesis that changes in crop shares are driven largely by Farm revenues and expenses that influence land rents. For the policy recommendations, increasing irrigation leads to an increase in croplands in particular cotton land share has a demand for irrigation in Alabama. Also increasing government payments for cropland will increase cotton and corn shares. Future research should consider expanding to other states in the Southeast such as Mississippi and Georgia to investigate the irrigation effects with land use and crop acreage decisions to get a clearer picture of the Southeast US. Also, the inclusion of variables like land prices, input and output prices, urban and rural population, metro and non-metro classification should provide additional information of land use and cropping decisions.

References

Adjei, E. (2020). Timberland Dynamics and Hunting Demand: Special Reference to Alabama.

Ahn, S., Plantinga, A. J., & Alig, R. J. (2000). Predicting future forestland area: a comparison of econometric approaches. *Forest Science, 46*(3), 363-376.

Alig, Ralph J. **(**1986). “Economic analysis of factors influencing forest acreage trends in the Southeast.”, *Forest Science* 32 (4): 19- 13.

Arellano M, Bond S (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev Econ Stud 58(2):277–297

Asseng S. (2013) Agriculture and Climate Change in the Southeast USA. In: Ingram K.T., Dow K., Carter L., Anderson J. (eds) Climate of the Southeast United States. NCA Regional Input Reports. Island Press, Washington, DC.

Barlowe, R. (1978). Land resource economics: the economics of real estate.

Bell, K. P., Boyle, K. J., Plantinga, A. J., Rubin, J., & Teisl, M. F. (2017). Overview of Empirical Methods . In Economics of Rural Land-Use Change (pp. 113-134). Routledge.

Bell, K. P., Boyle, K. J., & Rubin, J. (Eds.). (2006). Economics of rural land-use change. Ashgate Publishing, Ltd.

Blundell R, Bond S (1998) Initial conditions and moment restrictions in dynamic panel data models. J Econometrics 87(1):115–143

Bockstael, N.E. (1996). 'Modeling Economics and Ecology: The Importance of a Spatial Perspective.' *American Journal of Agricultural Economics* 78(5): 1168–80.

Chakir, R., & Lungarska, A. (2015). *Agricultural land rents in land use models: a spatial econometric analysis* (No. 718-2016-48744).

Chakir, R., & Le Gallo, J. (2013). Predicting land use allocation in France: A spatial panel data analysis. *Ecological Economics*, *92*, 114-125.

Claassen R, Tegene A (1999) Agricultural land use choice: A discrete choice approach. Agricultural and Resource Economics Review28: 26–36.

Davis K.F., C. Dalin, R. DeFries, J.N. Galloway, A.M. Leach, N.D. Mueller (2019). Sustainable pathways for meeting future food demand Encyclopedia Food Security Sustain., 3, pp. 14-20

Hardie, I. W., & Parks, P. J. (1997). Land use with heterogeneous land quality: an application of an area base model. *American Journal of Agricultural Economics*, 79(2): 299-310.

Hardie, I., P. Parks, P. Gottleib, and D. Wear. 2000. 'Responsiveness of Rural and Urban Land Uses to Land Rent Determinants in the U.S. South.' *Land Economics* 76(4): 659–73.

Kim, T. J., Wear, D. N., Coulston, J., & Li, R. (2018). Forest land use responses to wood product markets. *Forest policy and economics*, *93*, 45-52.

Kline, J.D., and R.J. Alig. (1999). 'Does Land Use Planning Slow the Conversion of Forest and Farm Lands?' *Growth and Change* 30(1): 3–22.

Lamb, M., Sorensen, R., Nuti, R. C., Butts, C. L., Faircloth, W. H., Eigenberg, D., & Rowland, D. L. (2011). Agronomic and economic effect of irrigation rate in corn produced in Georgia. *Crop Management*, *10*(1), 1-8.

Lambert, L. H., English, B. C., Clark, C. C., Lambert, D. M., Menard, R. J., Hellwinckel, C. M., & Papanicolaou, A. (2021). Local effects of climate change on row crop production and irrigation adoption. Climate Risk Management, 32, 100293.

Lichtenberg, E. (1989). Land quality, irrigation development, and cropping patterns in the northern high plains. *American Journal of Agricultural Economics*, *71*(1), 187-194.

Lubowski, R. N., Plantinga, A. J., & Stavins, R. N. (2008). What drives land-use change in the United States? A national analysis of landowner decisions. *Land Economics*, *84*(4), 529-550.

Lubowski, R. N., Bucholtz, S., Claassen, R., Roberts, M. J., Cooper, J. C., Gueorguieva, A., & Johansson, R. C. (2006). *Environmental effects of agricultural land-use change: The role of economics and policy* (No. 1477-2016-121102).

Magliocca, N., Kumar, M., & Handyside, C. T. (2020). Spatio-temporal patterns of irrigation adoption and water use in Alabama. In AGU Fall Meeting Abstracts (Vol. 2020, pp. GC041-0014).

Majumdar, I., Polyakov, M., Teeter, L. D., & Butler, B. J. (2009). Effect of population pressure on forest land use change in alabama: a nested logit approach. *International Journal of Ecological Economics and Statistics*, *14*(P09), 77-93.

Majumdar, I., Polyakov, M., & Teeter, L. D. (2007). Effect of Urbanization on the Forest Land Use Change in Alabama: A Discrete Choice Approach. Emerging issues in forest economics, 143.

Miao, R. (2013). Impact of ethanol plants on local land use change. *Agricultural and Resource Economics Review*, *42*(2), 291-309.

Miller, D.J., and A.J. Plantinga. (1999). 'Modeling Land Use Decisions with Aggregate Data.' *American Journal of Agricultural Economics* 81(1): 180–194.

Mu, J. E., Sleeter, B. M., Abatzoglou, J., & Antle, J. (2017). Climate impacts on agricultural land use in the United States: The role of socioeconomic scenarios. Climatic Change https://doi.org/10.1007/s10584-017-2033-x, 144, 329–345.

Mullen, J. D., Yu, Y., & Hoogenboom, G. (2009). Estimating the demand for irrigation water in a humid climate: A case study from the southeastern United States. *Agricultural Water Management*, *96*(10), 1421-1428.

Nagubadi, R. V., & Zhang, D. (2005). Determinants of timberland use by ownership and forest type in Alabama and Georgia. *Journal of Agricultural and Applied Economics*, *37*(1379-2016-112677), 173-186.

Napton, D. E., Auch, R. F., Headley, R., & Taylor, J. L. (2010). Land changes and their driving forces in the Southeastern United States. *Regional Environmental Change*, *10*(1), 37-53.

Newburn, D., Reed, S., Berck, P., & Merenlender, A. (2005). Economics and Land ‐ Use Change in Prioritizing Private Land Conservation. *Conservation Biology*, *19*(5), 1411-1420.

Nickell S (1981). Biases in dynamic models with fixed effects. *Econometrica 49*:1417–1426

Parks, P.J., and B.C. Murray. (1994). ‘Land Attributes and Land Allocation: Nonindustrial Forest Use in the Pacific Northwest.’ *Forest Science* 40(3): 558–575.

Pearson, K. (1901) On lines and planes of closest fit to systems of points in space. Philos. Mag. Ser.6,2, pp. 559-572.

Plantinga, A. J., Lubowski, R. N., & Stavins, R. N. (2002). The effects of potential land development on agricultural land prices. *Journal of Urban Economics*, *52*(3), 561-581.

Plantinga, A.J., Ahn, S., (2002). Efficient policies for environmental protection: an econometric analysis of incentives for land conversion and retention. Journal of Agricultural and Resource Economics 27 (1), 128–145.

Plantinga, A.J., T. Mauldin, and D.J. Miller. (1999). ‘An Econometric Analysis of the Costs of Sequestering Carbon in Forests.’ *American Journal of Agricultural Economics* 81(4): 812–24.

Plantinga, A.J., (1996). The effect of agricultural policies on land use and environmental quality. American Journal of Agricultural Economics 78 (4), 1082–1091.

Polyakov, M., & Zhang, D. (2008). Property tax policy and land-use change. *Land Economics*, *84*(3), 396-408.

Roodman D (2009a) How to do xtabond2: an introduction to difference and system GMM in Stata. Stata J 9(1):86–136

Roodman D (2009b) A note on the theme of too many instruments. Oxf Bull Econ Stat 71(1):135--158

Schaible, G., & Aillery, M. (2012). Water conservation in irrigated agriculture: Trends and challenges in the face of emerging demands. *USDA-ERS Economic Information Bulletin*, (99).

Sohl, T., & Sayler, K. (2008). Using the FORE-SCE model to project land-cover change in the southeastern United States. *ecological modelling*, *219*(1-2), 49-65.

Srivastava, P., Gupta, A. K., & Kalin, L. (2010). An ecologically-sustainable surface water withdrawal framework for cropland irrigation: A case study in Alabama. *Environmental management*, *46*(2), 302-313.

Stavins, R.N., Jaffe, A.B., (1990). Unintended impacts of public investments on private decisions: the depletion of forested wetlands. The American Economic Review 80 (3), 337–352.

USDA ERS (Economic Research Service). 2021. Irrigation and Water Use. Washington, DC: USDA Economic Research Service. https://www.ers.usda.gov/topics/farm-practices-management/irrigation-water-use/

Wu, J. (2008). Land use changes: Economic, social, and environmental impacts. *Choices*, *23*(4), 6-10.

Wu, J., and B.W. Brorsen. (1995). ‘The Impact of Government Programs and Land Characteristics on Cropping Patterns.’ *Canadian Journal of Agricultural Economics* 43(1): 87–104.

Wu, J., Segerson, K., (1995). The impact of policies and land characteristics on potential groundwater pollution in Wisconsin. American Journal of Agricultural Economics 77 (4), 1033–1047.

Zellner, A. 1962. An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. Journal of American Statistical Association. 57:348-368.

Zhou, N. (2010). *An econometric analysis of Alabama rural land values* (Doctoral dissertation).

2. CHAPTER 2

The Impact Analysis of Dodd-Frank Act on Financial Inclusion and Alternative Financial Services of Unbanked Rural households in the United States

Introduction

Financial inclusion has become a major topic of interest to the financial world for the last two-three decades. Dodd-Frank Wall Street Reform and Consumer Protection Act passed by the U.S. Congress on July 21, 2010, after the financial crisis to regulate the financial industry. This law made changes in response to the financial crisis that occurred in 2008 by regulating banks and protecting consumers from predatory and unfair financial practices affecting all federal financial regulatory agencies to prevent a future financial crisis (Geler, 2020). Dodd-Frank act created new federal agencies such as Consumer Financial Protection Bureau (CFPB) and other financial councils to achieve the purpose of financial regulation and consumer protection. This paper studies the characteristics of unbanked rural households on their intention to use formal financial services that lead to financial inclusion and alternative financial services (AFS) after the impact of the Dodd-Frank act.

Recent Federal Deposit Insurance Corporation (FDIC) survey in 2019; Household Use of Banking and Financial Services in the United States (U.S.) found that 5.4 percent of U.S. households nearly about 7.1 million were unbanked it means that no person in the household had a checking or savings account. The unbanked proportion of the U.S. in 2019, 5.4 percent was the lowest rate and in 2009, it was 7.6 percent. Based on the 2017 FDIC National Survey data on Unbanked and Underbanked Households, the unbanked rate was 6.5 percent, and it represents nearly 8.4 million U.S. households. Between 2017 and 2019, the unbanked rate dropped by 1.1 percentage points, compared to an increase of around 1.5 million banked households. Unbanked rate decline between 2017 and 2019 correlated with the improvement of the socio-economic conditions of U.S. households over this period.

Unbanked individuals are mostly like to be lower-income, younger, along with racial or ethnic minorities, have less formal education, disabled, and have varied monthly incomes compared to the general U.S. population (Congressional research service, 2019). When unbanked households do not have a bank account, they use financial products or services outside of the banking system termed as alternative financial services (AFS). Recent FDIC surveys show that unbanked households continued to use much higher AFS compared to banked households. Unbanked households that used AFS has decreased compared to declines in the use of both AFS transactions and credit in recent years as a proportion and AFS use among banked households also decreased over this period (FDIC survey, 2017).

There are many reasons rural households being unbanked and tend to use more AFS in the past decade. Survey of Consumer Finances (SCF) in 2010 reported that 11% of rural households did not have bank accounts and these unbanked households are often determined to turn to AFS such as check cashing, which is expensive and there were no fair rules to enforce the AFS providers. Also, a survey conducted by the Federal Reserve shows that rural households did not have a checking account (25%) because they did not like dealing with banks and 12 percent because the service charges were too high. Another reason is that customers couldn’t enroll in costly overdraft programs because of the CFPB will stop banks enrolling customers without their consent. Also, CFPB supervision and control larger AFS companies such as payday lenders and check cashers to prevent harmful practices that help families to avoid hidden fees and keep more money to help using AFS wisely.

Over the last three decades, more than half of all banks in the U.S. have closed and, these figures are even greater in rural areas due to the depopulation of rural counties; decreasing the need for technological advances for physical banks’ facilities, increased and adverse regulations of the Dodd-Frank Act (Wheelock, 2012). The closure of these banks harms small, local lenders by forcing on them one-size-fits-all financial parameters aimed at big Wall Street banks (Covington and Courtney, 2014). However, the most serious issue is that nearly 96 percent of them have been community banks of the bank closures, (FDIC, 2012).

This paper provides new evidence about the differences in financial inclusion in rural and household economies after the Dodd-Frank act and its effects. It uses data on household level data in the U.S., taken from the 2009, 2011 and 2019 FDIC National Surveys of Unbanked and Underbanked Households sponsored by the U.S. Census Bureau's Current Population Survey (CPS). This paper investigates the impact of the pre and post effect of the Dodd-Frank act on rural households' financial inclusion. For the purpose of this paper, financial inclusion defined as the level of which targeted at bringing unbanked rural households into the formal finance sector opposite to using alternative financial services (AFS). This paper relies on national household survey data collected by FDIC over the period 2009, 2011, and 2019 to corresponding demographic, economic, and financial data. Remaining sections of this paper is organized as follows. The next section provides a literature review of the importance of financial inclusion to US households and how it affects after the Dodd-Frank act. Third section provides the data description related to the econometric approach and the empirical model. The fourth section discusses the empirical results and discussions. Final section contain conclusions and limitations as plans for further improvement of the study.

Literature Review

Financial inclusion is a relatively new and dominant component in development economics, and it has gained attention over the last two decades. Financial inclusion has grown rapidly equal to education, healthcare, property rights, and infrastructure, which increase economic growth and reduce poverty (Karp and Nash-Stacey, 2015). According to the Federal Deposit Insurance Corporation (FDIC), financial inclusion is defined as bringing unbanked and underbanked consumers into the formal finance sector aiming various public and private efforts who access only alternative financial services (AFS). Also, financial inclusion defined as a measure of individuals and businesses have access to useful and reasonable financial products and services such as credit, transactions, payments, savings, and insurance that meet their financial needs (World Bank,2018). According to the FDIC, opening bank accounts can be one of the most important steps taken toward reaching the financial goals of unbanked households. Having a formal bank account provides benefits; financial safety, protection against error and fraud, easy access to funds and online purchases, proof of payments and bills you pay from any location, savings from check-cashing fees and overall financial peace of mind.

Financial inclusion linking the financial growth has provided literature that increasing financial sector will lead to address financial inclusion in the United States. The relationship between a country's domestic credit as a percent of GDP and the percent of adults that report having a formal bank account is positive and statistically significant and financial inclusion is more generally correlated with economic development (Demirgüç-Kunt and Klapper, 2013). There are individual countries that have been identified the determinants of being unbanked related to financial inclusion with variations over time. For example, the expansion of a Mexican savings institute, increased the average savings rate of affected low-income households and advanced to financial inclusion (Aportela, 1999). Burgess and Pande (2005) provide evidence of opening bank branches in unbanked rural areas (state-led expansion) in India associated with reduction of poverty of rural poor and led to financial inclusion. Using the same identification approach, Rhine, and Greene (2006) concluded that income, wealth, and education are important determinants of being unbanked. Osili and Paulson (2008) found that immigrants in the US who have more effective institutions in their countries than other immigrants are more likely to have a relationship with a bank and use more formal banking services.

Previous studies suggest that financial inclusion provide economic benefits. Ruiz (2013) showed that formal financial service providers help to cover unexpected expenses of households who have generated income from savings accounts. Financial inclusion also provides other benefits at the individual level as found in various literature. Beck, Demirgüç-Kunt and Levine (2007) found lower inequality linked to economy in wide level, female empowerment (Ashraf, Karlan, and Yin 2010), greater entrepreneurship with gain access to financial services (Banerjee et al. 2010; Karlan and Zinman 2010; Demirgüç-Kunt and Klapper 2013), larger investment in education and businesses (Brune, Giné, Goldberg, and Yang 2011) and better health (Dupas and Robinson 2013).

There are recent studies have focused on financial inclusion such as Celerier and Matray (2019) examined the wealth accumulation along with the financial well-being related to the financial inclusion. They found that the U.S. branching deregulation increased the supply of bank branches and financial inclusion. Furthermore, they find that financial inclusion positively affects the wealth accumulation of low-income household, and households who is having a bank account allow not only to accumulate liquid assets, but also permanent assets. More generally, their results indicate that the increase of wealth accumulation by low-income households increases the access to a bank account. Dunham (2019) examined the relationship between sociodemographic characteristics and mortgage lending variables when AFS providers such as check cashing outlets are more widespread than banks on the presence of census regions in southeastern Pennsylvania and found that the availability of more check-cashing outlets than banks is predicted because of lower median household income, lower percentage of residents aged 65 or above, higher percentage of Black and Latin residents among other factors.

There are number of barriers to financial inclusion regardless of the many benefits. Ashraf, Karlan, and Yin (2006) showed that there is a potential barrier with upfront cost and other fees associated with opening a bank account, minimum balance requirement to keep up the account, requirement of necessary documents and proof of identity, and opportunity costs associated with opening an account and cost linked with traveling to a bank branch. Other potential barriers such as distrust of banks and lack of financial capability related to financial inclusion is not entirely clear (Karlan, Ratan and Zinman 2014; Fernandes, Lynch and Netemeyer 2014).

There are major barriers as the reasons of unbanked households that prevent entering the formal banking sector. Hayashi (2013) used FDIC and Board of Governors of the Federal Reserve System surveys to find main reasons of unbanked households are not having bank accounts. He specified that the main reason of consumers that they do not use banks due to the high maintenance cost of an account because of their low, unstable income and banks' high fees and other reasons such as negative perceptions or experiences with banks; do not meet banks' qualification requirements; the physical accessibility of banks, linked with locations and hours; consumers' privacy, relates to their preference, and the characteristics of bank accounts and related payment services that do not meet the needs of certain groups of unbanked consumers.

Rhine and Greene (2013) explored the reasons of becoming unbanked in the United States and found that families more likely to become unbanked when there is a significant decline in the loss of employment, or loss of health insurance coverage and family income. Campbell, Martinez-Jerez and Tufano (2012) show that unintentional bank account closures are more frequent when U.S. counties have lower wealth, lower education and higher unemployment. Also, higher rates of involuntary bank account closures occur due to the access to payday lending.

FDIC's National Survey of Unbanked Households shows similar findings to other studies in recent years. Barr (2002), Caskey (2002), and Hogarth et al. (2005) used 1990s and early 2000s data to find connections between the sociodemographic characteristics and unbanked rate. Sherraden (2010) finds that individuals who unmarried, less educated, minorities, young, and less wealthy likely to be without the use of checking, or savings accounts. Similarly, unbanked households show attributes such as low income, young, immigrants, less education, and female head of the household (Beard, 2010). Also, age, race/ethnicity, education, employment, home ownership, and spending habits are significantly associated with households who are not engaged in typical banking (Vermilyea and Wilcox, 2002).

Berry (2004) analyzed the reasons why people choose to be unbanked by collecting data from low income families in Los Angeles, Chicago, and Washington DC and suggests that immigrants more likely use AFS for their needs. Also, in New York, thousands of low income families choose to stay unbanked because of excessive banking fees and turn to non-traditional financing services, partly because they lack the educational awareness to choose formal banking services (NY Neighborhood Financial Services Study, 2008). Also, educational background, legal status, income, longevity of residence in the United States and English language proficiency are found to be factors that impact the quality of life and remain unbanked or underbanked in immigrant communities. Smith et al. (2008) investigated and found that AFS providers are located in areas that lack of access to traditional banks. Bradley et al. (2009) found that Nonbank AFS institutions providing services that include money transmission, car title lending, pawn shops and rent-to-own stores.

According to a recent report of the Board of governors of the federal reserve system, number of bank branches in the rural communities are more or same in 2017 compared to 2012, and over 40 percent of rural counties lost bank branches during 2012-2017 period.

Rural communities who are poorer residents, less likely to finish high school or college degree, or being African American are deeply affected in these 5 years compared to less affected communities with the experience of significant declines in the availability of bank branches in the rural counties over that period (Board of governors of the Federal reserve system, 2019).

The Dodd-Frank Act addresses many issues that policymakers reason contributed to the financial crisis in 2008 (Le, 2017). However, the Dodd-Frank Act has unreasonably impacted on community banks because these community banks depend on limited resources of financial support and do not benefit from economies of scale. By subjecting these community banks of all asset sizes to fixed compliance costs and to many of the same regulations, it has become economically tough for community banks to provide access to basic banking services to rural communities without making profits that cover increased compliance costs (Schorgl, 2018).

Data and Methodology

This paper relies on the survey data from the FDIC National household survey from the years of 2009, 2011 and 2019. The research question of this paper is “How has the unbanked rural households’ opening bank accounts and using alternative financial services affected after the Dodd-Frank act of 2010?” This household survey data provide information on like to open a bank account, using AFS credit services and AFS transaction services, households’ demographic and socioeconomic information. The number of observations of unbanked households extend 3033 (4.2%) from 2009, 3219 (4.5%) from 2011 and 1611 (2.3%) from 2019 samples. Table 2.1 provides a description of all the covariates used in the study.

Table 2.1. Description of all variables of Unbanked rural and urban households included in the model

Variable	Variable description
Like to open a bank account	1 = Not at all likely to open a bank account
	2 = Not very likely to open a bank account
	3 = Somewhat likely to open a bank account
	4 = Very likely to open a bank account
AFS Credit use	1 = Household use AFS credit services
	0 = Household do not use AFS credit services
AFS Transaction use	1 = Household use AFS transaction services
	0 = Household do not use AFS transaction services
Age	Household respondent's age
Number of persons in the household	Number of people who live in the household
Rural	1 = Household live in the rural/non-metropolitan area
Family income less than $15k	1 = Family income of the household is less than $15,000
Highschool diploma	1 = Education level of the household is high school diploma
College degree	1 = Education level of the household is a college degree
Employed	1 = Household is employed
US born / foreign born citizen	1 = Household is a US born or foreign born citizen in US
Race / Ethnicity	1 = Black
	2 = Hispanic
	3 = Asian
	4 = White
	5 = Other
Married	1 = Marital status of the household is married
Female headed family	1 = Household type identified as a female head
Homeowner	1 = Respondent is identified as a home owner
Previously banked	1 = Household had used bank services or an account earlier
Bank branches density	Number of bank branches per 100,000 people in a county

Using these variables, I analyze the impact of the Dodd-Frank act on financial inclusion of unbanked rural households constructing empirical model using two periods in Changes in Changes (CIC) method. In this study, only two groups and three periods are considered (2009 as the pre period and 2011 and 2019 as post periods). Athey and Imbens (2006) suggest the CIC model as a non-parametrical approach and an alternative to the Difference-in-difference (DID) model. The CIC method relaxes more restrictive assumptions and treats groups and time periods asymmetrically and also relaxes the parallel trend assumption holding the rank preservation assumption (De Alwis, 2020). This method uses the entire "before" and "after" outcome distributions of the treated group to non-parametrically estimate the change occurred in the treated group over time and recovers the whole distribution of the counterfactual outcome. This estimation is relatively straightforward in the absence of covariates. Melly and Santangelo (2015) extend it to the case where the identifying assumptions hold conditional on covariates. This semi parametric estimator incorporate covariates in the Athey and Imbens (2006)'s procedure to obtain unconditional estimates. Chernozhukov et al. (2013b) consider identification of the conditional Average Treatment Effect (ATE) and Quantile Treatment Effect (QTE) for nonseparable panel data models under a time homogeneity condition. D'Haultfoeuille et al. (2015) present identification of nonseparable models using repeated cross sections.

Using the empirical design of the difference in-differences approach to examine the differential trends there are only two groups (g = u, r) observed in two time periods (t = 1, 2); this situation is often represented by a 2×2 matrix experience outcomes in y. In this study, y_1 is the outcome in the presence of the Dodd-Frank effect as the treatment and y_0 is the outcome in the absence of the Dodd-Frank effect. I examine the Likelihood of an opening a bank account; AFS credit use and AFS transaction use as the outcomes of this study. Group u is the control group, and the group r is the treatment group. t = 1 taken as the pre-treatment time period, and t = 2 is the post-treatment time period. The treatment is observed only if g = r and t = 2. In our two-period difference-in-differences context, time t = 1 corresponds to 2009, the year before the Dodd-Frank act implemented, and time t = 2 corresponds to 2011 and 2019, the last sample year. Group g = 2 is rural households that affected after the Dodd-Frank act by the years 2011 and 2019.

This study adopts the convention of two categories of metropolitan (or metro) and nonmetropolitan (or nonmetro) areas to define the urban and rural households classified by United States Department of Agriculture-Economic Research Service (USDA-ERS) . Metropolitan Statistical Areas (MSAs), each of which includes an urban area with population of at least 50,000 people and non-metropolitan statistical areas defined by populations with fewer than 50,000 people. This study employs these definitions as the availability of data to prior years. In the most basic model, the average treatment effect on the treated in the effect of act can be written as equation (1):

$$\tau DID = E[y_{r2}^{1}] - E[y_{r2}^{0}] = E[y_{r2}] - E[y_{r1}] - (E[y_{u2}] - E[y_{u1}]) \tag{1}$$

The basic empirical model can be expressed as the model is given at individual household-level *i* in time period *t*

$$Y_{it} = \mathrm{f}(T_t, \mathrm{Z}, X_t) \tag{2}$$

where Y_{it} is the outcome variable; T_t and given binary time period *t*; Z is variable representing treatment group (rural and urban households) and X_t is the vector of control variables (covariates) representing the households' characteristics at time period *t* shown in equation (2). A regression equation to this model allows us to control for observable differences in the distribution of characteristics of the treatment and control groups (Beatty and Shimshack, 2011). This regression model is parameterized following the difference-in-differences literature and can be written as:

$$Y_{it} = \beta_0 + \beta_1 Time + \beta_2 Treatment + \beta_3 (Time \times Treatment) + \beta_4 X_{it} + \varepsilon_{it} \quad (3)$$

where Y_{it} is like to open a bank account, use of alternative financial services as credit (payday, pawn shop, refund anticipation and title loan, rent-to-own service etc.) and use of alternative financial services as transactions (Check cashing, money order) by an individual household *i* at time period t. The vector X_i contains the set of demographic and socio-economic characteristics age group, education, employment status, family income, Spanish only language spoken, nativity, race, household type, home ownership and previously banked status of households. ε represents the standard idiosyncratic disturbance error term.

In general, quantile regression model estimates the conditional distribution outcome at a different point which is a linear function of the covariates (Koenker and Bassett 1978; Chernozhukov and Hansen 2006) compared to the linear regression model that estimate conditional expectations of outcome (De Alwis, 2020). In the spirit of the DID estimation, we adopt the change-in-changes method outlined in Athey and Imbens to estimate the impact of the Dodd-frank act on the distribution of the outcome variables. The treatment effect is identified as

$$\gamma^{CIC} = F^{-1}_{Y_{1|D=1}}(\gamma) - F^{-1}_{Y_{0|D=1}}(\gamma) \quad (4)$$

where $F_{Y_{1|D=1}}(.)$ represents the distribution of treated potential outcomes for the treated group and $F_{Y_{0|D=1}}(.)$ is the distribution of untreated potential outcomes for the treated group. $F_{Y_{1|D=1}}(y)$ is identified directly because we observe the distribution of the treated outcome for the treated. $F_{Y_{0|D=1}}(.)$ cannot be directly identified from the data. Therefore, the estimator uses as a proxy for the change in the outcome variables that would have occurred for level y in the treatment group (unbanked rural) in the absence of the policy introduction, namely, the change in the outcome that did occur in the control group at the quantile of the control group that corresponds to that at the same level y.

The Change-in-Changes approach (CIC) use as our estimation method to isolate the quantile treatment effect. I use CIC method to estimate how the Dodd-frank act affected the outcome distribution of financial inclusion indicators of unbanked rural households. I estimate the quantile treatment effects for the year after the policy using the Stata16 and R softwares. The estimated quantile treatment effects (QTEs) allow us to assess the impact of the policy on the lower, middle and upper parts of the outcome distribution.

Results and Discussion

Table 2.2 present the summary statistics for the variables of rural and urban unbanked households and provide the information on the demographic and county characteristics of rural and urban households for the years 2009, 2011 and 2019 to measure the degree of the financial inclusion before and after the Dodd-Frank policy act of 2010. The first column reports the results for the whole unbanked population. Columns 2 and 3 describe the results for unbanked households that live in rural and urban regions according to the metropolitan status. Compared with households who live in rural areas, the households live in urban areas tend to be younger (44 years) and more likely to have members in the household. Compared to rural households, urban households earn more money as the low value of family income less than $15k on average. Bank branches density per 100,000 people is close to 32 for the rural households comparing the bank branches density of 46.06 for urban sample.

Table 2.2. Summary statistics of total unbanked, rural and urban households

Household Variables	**Total unbanked** (N = 7801) Mean (SD)	**Unbanked rural** (Treated, n=1767) Mean (SD)	**Unbanked urban** (Control, n=6034) Mean (SD)
Outcome variables			
Like to open a bank account	1.96 (1.06)	1.86 (1.02)	1.99 (1.07)
AFS Credit use	0.28 (0.45)	0.30 (0.46)	0.28 (0.45)
AFS Transaction use	0.66 (0.47)	0.65 (0.47)	0.66 (0.47)
Demographic variables			
Age of the household	44.47 (16.22)	45.86 (16.68)	44.06 (16.07)
Number of persons in the household	2.58 (1.70)	2.47 (1.57)	2.62 (1.73)
Family income less than $15000	0.49 (0.50)	0.56 (0.50)	0.47 (0.50)
Highschool diploma	0.38 (0.48)	0.40 (0.49)	0.37 (0.48)
College degree	0.05 (0.21)	0.02 (0.15)	0.05 (0.22)
Employed	0.41 (0.49)	0.36 (0.48)	0.42 (0.49)
U.S. born / foreign born citizen	0.84 (0.37)	0.94 (0.23)	0.80 (0.40)
Black	0.34 (0.47)	0.37 (0.48)	0.33 (0.47)
Hispanic	0.25 (0.43)	0.29 (0.09)	0.30 (0.46)
Asian	0.02 (0.13)	0.01 (0.11)	0.02 (0.14)
White	0.38 (0.48)	0.56 (0.49)	0.32 (0.47)
Other	0.04 (0.18)	0.27 (0.08)	0.15 (0.02)
Married	0.22 (0.42)	0.23 (0.42)	0.22 (0.41)
Female headed family	0.28 (0.45)	0.29 (0.45)	0.28 (0.45)
Homeowner	0.25 (0.43)	0.39 (0.49)	0.21 (0.40)
Previously banked	0.50 (0.50)	0.54 (0.50)	0.48 (0.50)
Bank branches density (per 100,000 people)	35.06 (20.65)	31.86 (20.20)	46.01 (18.30)

Also, figures A.1,A.2 and A.3 show the percentages of unbanked households' likelihood of opening a bank account, AFS credit uses and AFS transaction uses for 2009, 2011 and 2019. These graphs present that there is a gradual increment in all three factors as opening bank accounts, AFS credit use and AFS transaction uses for this 2009-2019 time period. Table

2.3 show the means of outcome variables, demographic and county variables for treated group (rural) and control group (urban) of unbanked households for 2009, 2011 and 2019.

Table 2.3. Sample means of Unbanked households for 2009, 2011 and 2019

Household Variables	Treated (Rural)			Control (Urban)		
	2009	2011	2019	2009	2011	2019
Outcome variables						
Likelihood of opening a bank account	1.85	1.95	1.69	2.02	2.09	1.72
AFS Credit use	0.30	0.38	0.17	0.30	0.34	0.12
AFS Transaction use	0.68	0.73	0.49	0.68	0.71	0.54
Demographic variables						
Age of the household	44.85	44.59	50.11	42.37	43.66	48.06
Number of persons in the household	2.53	2.55	2.23	2.73	2.65	2.34
Family income less than $15000	0.53	0.58	0.59	0.45	0.48	0.47
Highschool diploma	0.41	0.41	0.36	0.37	0.37	0.38
College degree	0.01	0.02	0.05	0.05	0.05	0.06
Employed	0.41	0.34	0.33	0.45	0.40	0.41
U.S. born / foreign born citizen	0.94	0.94	0.96	0.79	0.80	0.83
Black	0.21	0.22	0.24	0.34	0.33	0.33
Hispanic	0.11	0.13	0.09	0.30	0.30	0.29
Asian	0.01	0.01	0.01	0.02	0.02	0.02
White	0.56	0.56	0.58	0.31	0.33	0.34
Other	0.11	0.09	0.09	0.02	0.02	0.02
Married	0.25	0.22	0.20	0.24	0.22	0.19
Female headed family	0.28	0.31	0.25	0.29	0.29	0.24
Homeowner	0.42	0.37	0.37	0.20	0.21	0.22
Previously banked	0.57	0.52	0.52	0.51	0.45	0.49
Bank branches density (per 100,000 people)	32.49	29.95	24.63	49.26	40.99	36.83

There are three key outcome variables in our empirical analysis. These are likelihood of opening bank account, AFS credit use, and AFS transaction use of unbanked urban and rural households. The Changes in changes (CIC) model specification included rural, family income less than $15,000, having a high school diploma or a college degree for education, employed or not, U.S. born or foreign born citizen to control nativity, Black or white household in race, married or female headed family to control household type, home ownership and previously banked status as dummy variables and age, number of persons in the household and bank branches density per 100,000 people as covariates to control the outcome variables; likelihood of opening a bank account, AFS credit use and AFS transaction use after the treatment effect across treated unbanked rural groups.

CIC method proposed by Melly and Santangelo (2015) allows only two periods in the model with covariates and according to that in this model it reveals only average treatment effect of financial inclusion in the period after the implementation of the act. I estimated the quantile treatment effects with and without covariates for 2009-2011 and 2009-2019 time periods separately. Table 2.4 and Table 2.5 show the quantile treatment effects for treated (QTET) for 2011. Table 2.4 shows the quantile treatment effects of unbanked rural households using CIC model without covariates for 2011.

Table 2.4. Quantile treatment effects without covariates of Unbanked Rural households for 2011

Year Quantile	2011					
	Likelihood of Opening a bank account		AFS Credit use		AFS Transaction use	
	QTE	SE	QTE	SE	QTE	SE
0.1	-1.00	0.00	-1.00	0.00	-1.00	0.00
0.2	-1.00	0.00	-1.00	0.00	-1.00	0.32
0.3	-1.00	0.00	-1.00	0.00	0.00	0.00
0.4	-1.00	0.00	-1.00	0.00	0.00	0.00
0.5	0.00	0.53	-1.00	0.00	0.00	0.00
0.6	-1.00	0.00	0.00	0.52	0.00	0.00
0.7	0.00	0.70	0.00	0.00	0.00	0.00
0.8	-1.00	0.00	0.00	0.00	0.00	0.00
0.9	0.00	0.48	0.00	0.00	0.00	0.00

Notes: ***, **, *, stand for significance at 1%, 5% and 10% respectively.

The likelihood of opening a bank account quantile treatment effect shows that it increases negative effect of -1.0 to 0 in 0.5 quantile and 0.7 and 0.9 quantiles keep at 0 without any positive effect. In addition, the quantile treatment effect for AFS credit use shows that the using alternative financial credit services increase the negative effect to 0 after 0.5 quantile after the policy and AFS transaction use increases the negative effect to 0 after 0.2 quantile and not significant. According to the average quantile treatment effects of unbanked rural households show -0.80, -0.60 and -0.26 on likelihood of opening bank accounts, AFS credit use and transaction use for 2011 respectively without controlling covariates. These values mean that opening a bank account and using AFS credit and transaction services only show negative effects and not significant after the Dodd-Frank act in 2011. Appendix table A.1 and Figure 2.1 clearly depict these observations.

Table 2.5. Quantile treatment effects with covariates of Unbanked Rural households for 2011

	2011					
Year Quantile	Likelihood of Opening a bank account		AFS Credit use		AFS Transaction use	
	QTE	SE	QTE	SE	QTE	SE
0.1	0.137	0.104	-1.00	0.52	0.00	0.00
0.2	-0.053	0.108	-1.00	0.52	0.00	0.00
0.3	-0.243	0.178	-1.00	0.52	0.00	0.00
0.4	-0.467	0.233	-1.00	0.00	0.00	0.00
0.5	0.333**	0.256	-1.00	0.00	0.00	0.00
0.6	0.085**	0.114	0.00	0.52	0.00	0.00
0.7	0.661*	0.397	0.00	0.00	-1.00	0.42
0.8	0.000	0.064	0.00	0.00	0.00	0.52
0.9	0.000	0.390	0.00	0.00	0.00	0.00

Note: ***, **, *, stand for significance at 1%, 5% and 10% respectively.

The likelihood of opening a bank account quantile treatment effect shows that 0.1 quantile shows positive effects, and 0.2 to 0.4 quantiles show negative effects, and 0.5 to 0.7 quantiles show significant and positive effect to open a bank account. So, the average treatment effect is positive and close to 0.2. In addition, the quantile treatment effect for AFS credit use shows that using alternative financial credit services have a negative effect and increases after 0.5. AFS transaction use showed only a negative effect for 0.7 quantile. Average treatment effects for AFS credit use and transaction use are negative -0.54 and -0.06 for unbanked rural households in 2011. According to the quantile treatment effects, unbanked rural households had a positive significant effect on opening bank accounts after the Dodd-Frank act in 2011 with controlling demographic and county variables. Figure 2.1 clearly depicts these observations.

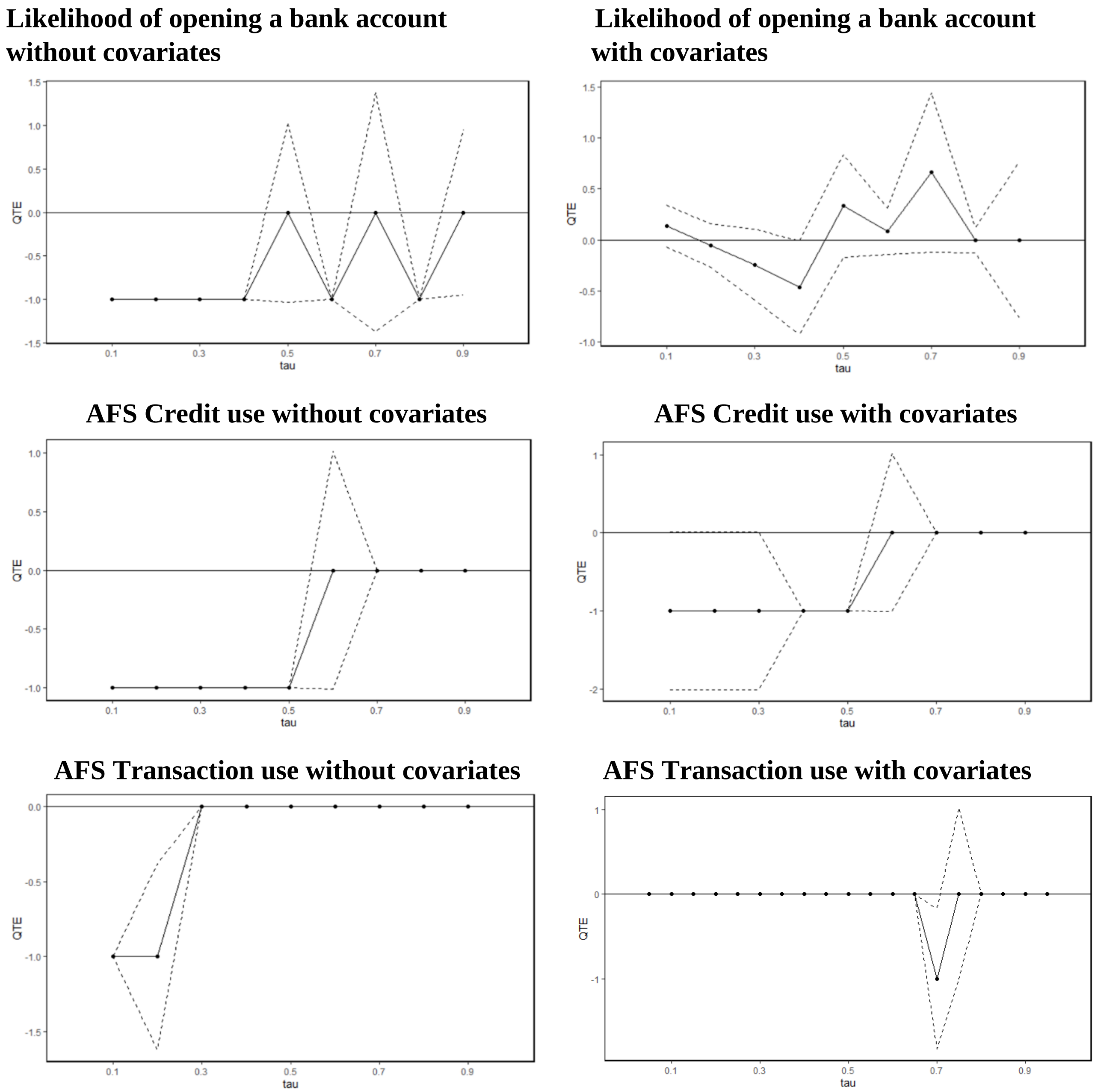

Figure 2.1. Quantile Treatment Effects for Unbanked Rural households' likelihood of opening a bank account, AFS credit use and AFS transaction use in 2011 after the Dodd-Frank act.

Notes: The left side figures provide estimates of the QTE for unbanked rural households using the no-covariates of the CIC method. The right side figures provide QTE estimates after conditioning on covariates. Standard errors are computed using the bootstrap with 100 iterations.

Table 2.6 and Table 2.7 show the quantile treatment effects for treated (QTET) for 2019. Table 2.6 shows the quantile treatment effects of unbanked rural households using the CIC model without covariates for 2019.

Table 2.6. Quantile treatment effects of Unbanked Rural households without covariates for 2019

Year Quantile	2019					
	Likelihood of Opening a bank account		AFS Credit use		AFS Transaction use	
	QTE	SE	QTE	SE	QTE	SE
0.1	0.00	0.00	0.00	0.00	0.00	0.00
0.2	0.00	0.00	0.00	0.00	0.00	0.00
0.3	0.00	0.00	0.00	0.00	0.00	0.48
0.4	0.00	0.00	0.00	0.00	-1.00	0.00
0.5	0.00	0.52	0.00	0.00	0.00	0.42
0.6	-1.00	0.53	0.00	0.00	0.00	0.00
0.7	0.00	0.32	-1.00	0.53	0.00	0.00
0.8	0.00	0.32	-1.00	0.42	0.00	0.00
0.9	0.00	0.57	0.00	0.00	0.00	0.00

Note: ***, **, *, stand for significance at 1%, 5% and 10% respectively.

In the table 2.6, likelihood of opening a bank account quantile treatment effect shows negative effect to opening a bank account at 0.6 quantile and while other quantiles stay 0. Average treatment effect is negative for opening a bank account which is -0.18 without covariates. The quantile treatment effects for AFS credit and transaction uses show negative effects in 0.7 and 0.8 for credit 0.4 quantile for transaction respectively. Overall treatment effects for AFS credit and transaction use are -0.14 and -0.17 in 2019 after the policy and not significant (Appendix Table A1).

Table 2.7. Quantile treatment effects of Unbanked Rural households with covariates for 2019

Year Quantile	2019					
	Likelihood of Opening a bank account		AFS Credit use		AFS Transaction use	
	QTE	SE	QTE	SE	QTE	SE
0.1	0.00	0.00	0.00	0.00	0.00	0.00
0.2	0.00	0.00	0.00	0.00	0.00	0.037
0.3	0.00	0.00	0.00	0.00	0.00	0.349
0.4	0.00	0.00	0.00	0.00	-1.00	0.00
0.5	0.00	0.00	0.00	0.00	0.00	0.483
0.6	0.00	0.57	-1.00	0.15	0.00	0.00
0.7	0.44**	0.45	-1.00	0.48	0.00	0.00
0.8	0.82**	0.49	0.00	0.00	0.00	0.00
0.9	0.00	0.37	0.00	0.00	0.00	0.00

Note: ***, **, *, stand for significance at 1%, 5% and 10% respectively.

In 2019, the likelihood of opening a bank account shows positive and significant effects only in 0.6 and 0.7 quantiles to open a bank account. In 0.6 quantile shows 0.44 and 0.7 quantile shows 0.82 treatment effects on likelihood of opening a bank account. So, the average treatment effect is positive (0.11) means that the effects are stronger at the lower tail of the distribution, and it makes a positive effect on like to open a bank account which is a good outcome after the policy. In addition, the quantile treatment effect for AFS credit use shows negative effects in 0.6 and 0.7 quantiles. AFS transaction use showed only a negative effect for 0.4 quantile. AFS credit use and transaction use did not show any positive and significant quantile treatment effects for 2019 and average treatment effects are -0.14 and -0.18 respectively. Figure 2.2 depicts these observations.

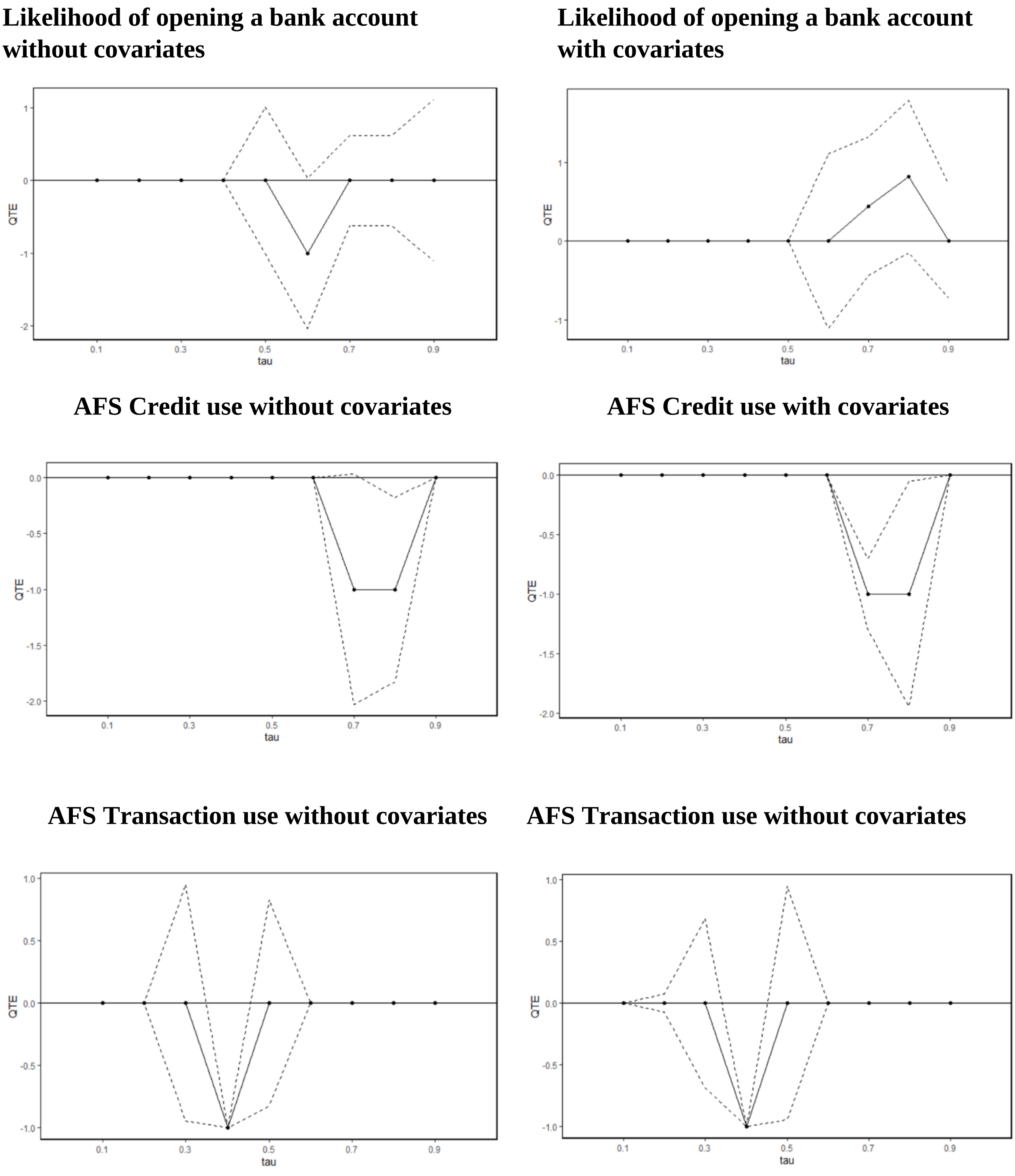

Figure 2.2. Quantile Treatment Effects for Unbanked Rural households' likelihood of opening a bank account, AFS credit use and AFS transaction use in 2019 after the Dodd-Frank act.

Notes: The left side figures provide estimates of the QTE for unbanked rural households using the no-covariates of the CIC method. The right side figures provide QTE estimates after conditioning on covariates. Standard errors are computed using the bootstrap with 100 iterations.

Table 2.8. Average Treatment effect (ATE) for the outcomes of unbanked rural households with covariates

Outcome variable	2011			2019		
	Average Effect	S.E.	Obs.	Average Effect	S.E.	Obs.
Like to open a Bank account	0.19**	0.10	6190	0.11**	0.10	4619
AFS Credit services use	-0.54	0.14	6190	-0.14	0.04	4619
AFS Transaction services use	-0.06	0.23	6190	-0.18	0.048	4619

Note: ***, **, *, stand for significance at 1%, 5% and 10% respectively.

Table 2.8 present the coefficient estimates of opening bank account, AFS credit and transaction uses of unbanked rural households controlling covariates in 2011 and 2019. On average, unbanked rural households are positive in opening bank accounts with a significant effect in 2011. Compared with households live in urban areas, those residing in rural areas have decreased effect using AFS credit and transaction services but not significant. Also, column 4 in Table 2.8 shows the average treatment effect of opening a bank account and it has decreased with a significant effect in 2019. Yet, the decrease of using AFS credit and transaction services in 2011 and 2019 are not statistically significant. According to the CIC estimates in 2011 and 2019, unbanked rural household’s likelihood of opening bank account has a positive and increasing effect in short term and long term after the Dodd-frank act compared to unbanked urban households with the control of demographic and county characteristics although the effect in 2019 has decreased compared to effect in 2011. So, the Dodd-Frank act impact decreases in time with the control of bank branches. AFS credit use and transaction use by unbanked rural households did not show significant effects in both short term and long term after the policy.

Robustness Checks

The empirical results are presented in Tables 2 9 and 2.10 and these tables display the odd ratio values from the ordered logistic regression models for opening bank accounts of unbanked households.

Table 2.9. Ordered Logistic Regression Results for Unbanked households for 2009, 2011 and 2019

Dependent Variable	**Likelihood of opening a bank account**	
	Odds Ratio	**SE**
Household control variables		
Age *(Years)*	0.969***	0.002
Number of persons in the household	1.009	0.018
Family income less than $15000 *(Yes or no)*	0.991	0.049
Highschool diploma *(Yes or no)*	1.048	0.051
College degree *(Yes or no)*	1.319**	0.152
Employed *(Yes or no)*	1.189***	0.060
U.S. born or foreign born citizen *(Yes or no)*	0.978	0.079
Race / Ethnicity (*White - Reference category*)		
Black	1.317***	0.078
Hispanic	1.013	0.076
Asian	1.884***	0.340
Other	1.037	0.131
Married *(Yes or no)*	1.316***	0.095
Female headed family *(Yes or no)*	1.317***	0.084
Homeowner *(Yes or no)*	0.961	0.056
Previously banked/ Had a bank account *(Yes or no)*	2.864***	0.144
County control variables		
Bank density *(Number of branches per 100,000 people)*	0.998	0.002
Interactions		

Rural × 2009	0.774***	0.072
Rural × 2011	0.789***	0.069
Rural × 2019	0.812	0.114
Year Dummies		
2011	Yes	
2019	Yes	
Number of observations	6845	
LR $chi^2(17)$	1125.38	
Prob > chi^2	0.0000	
Pseudo R^2	0.5653	
Log likelihood	-8051.50	

Notes: ***, **, *, stand for significance at 1%, 5% and 10% respectively.

Table 2.9 shows the relationship between household and county control variables that measure the extent of financial inclusion as the likelihood of opening a bank account with all unbanked households. The results indicate that age, college degree, employed, female headed family, black and married households and previously banked show significant odd ratios associated with household being unbanked. The proportional odds ratio for age of unbanked household explain that one unit increase of a variable, the odds of very likely to opening a bank account versus the combined somewhat likely, not very likely and not at all likely categories. Opening a bank account by unbanked household is 0.03 less likely and decreases with the age when other variables are held constant in the model. Having a college degree, the odds on very likely to open a bank account is 32% greater and the proportional odds ratio being employed 1.19 times greater for unbanked households. The logistic analysis indicates that Black households are 1.317 times more Asian households are 1.884 times more likely to have access to banking services as compared to the white households. The odds ratios for being married and female headed family show positive significant effects on opening a bank account as 1.316 and 1.317 are the greater odds for unbanked. Also, previously banked households 2.864 times very likely to open a bank account compared to households who are not previously banked. Unbanked rural households 0.774 times less likely and 0.789 times less likely to open a bank account for being compared to being unbanked urban households in 2009 and 2011 respectively. In 2019, unbanked rural households 0.812 times less likely to open bank accounts and it’s not significant value. Year dummies for 2011 and 2019 show effects on the likelihood of opening a bank account.

Table 2.10 shows the relationship between household and county control variables that measure the extent of financial inclusion as the likelihood of opening a bank account with the unbanked rural and unbanked urban households.

Table 2.10. Ordered Logistic Regression Results for Unbanked rural and urban households from 2009, 2011 and 2019

Dependent Variable	**Likelihood of opening a bank account**	
	Unbanked rural	**Unbanked urban**
Household control variables		
Age	0.964***	0.971***
Number of persons in the household	1.034	1.002
Family income less than $15000	1.017	1.002
Highschool diploma	1.047	1.045
College degree	2.639***	1.202
Employed	1.130	1.185***
U.S. born or foreign born citizen	1.090	0.965
Race / Ethnicity		
White (Reference category)		
Black	0.857	1.376***
Hispanic	1.025	1.018
Asian	2.159	1.730***
Other	1.126	0.903
Married	1.197	1.365***

Female headed family	1.402**	1.310***
Homeowner	1.005	0.963
Previously banked/ Had a bank account	3.068***	2.676***
County control variables		
Bank density	0.997	0.999
Year Dummies	Yes	Yes
Number of observations	1580	5265
LR $chi^2(15)$	307.32	725.01
Prob > chi^2	0.0000	0.0000
Pseudo R^2	0.4808	0.4541
Log likelihood	-1747.95	-6339.43

Notes: ***, **, *, stand for significance at 1%, 5% and 10% respectively.

The results indicate that age, having a college degree, female headed family, previously banked show positive and significant odd ratios associated with unbanked rural households. Also, age, employed, black, married and being previously unbanked urban households show significant associations with the opening a bank account. The proportional odds ratio for age of unbanked rural household explains that 0.036 less likely to open a bank account and decrease with the age. Likewise, unbanked urban households are opening accounts 0.971 times lower, given the other variables are held constant. The proportional odds ratio for having a college degree the odds on very likely to open a bank account are 2.639 times greater for unbanked rural households. The odds ratios for female headed family show positive significant effect on opening a bank account and 1.402 for unbanked rural households very likely to open a bank account compared to other categories**.** Also, previously banked unbanked rural households 3.068 times very likely to open a bank account compared to households who are not previously banked.

Table 2.11. Binomial Logistic Regression Results for Unbanked rural and urban households from 2009, 2011 and 2019

Dependent Variable	AFS Credit use	AFS Transactions use
Household control variables		
Age	0.981***	0.989***
	(0.002)	(0.002)
Number of persons in the household	1.112***	1.044*
	(0.028)	(0.025)
Family income less than $15000	1.021	0.938
	(0.073)	(0.061)
Highschool diploma	1.046	1.020
	(0.073)	(0.066)
College degree	0.567***	0.509***
	(0.101)	(0.067)
Employed	1.140*	1.381***
	(0.082)	(0.093)
U.S. born or foreign born citizen	3.157***	1.191*
	(0.408)	(0.115)
Race / Ethnicity		
Black	0.651***	1.080
	(0.052)	(0.082)
Hispanic	0.586***	1.054
	(0.059)	(0.097)
Asian	0.267***	0.355***
	(0.104)	(0.082)
Other	1.012	1.149
	(0.207)	(0.233)
Married	1.255**	1.215**
	(0.134)	(0.116)
Female headed family	1.062	1.100
	(0.097)	(0.093)
Homeowner	0.666***	0.740***
	(0.060)	(0.057)
Previously banked/ Had a bank account	2.647***	2.148***
	(0.188)	(0.138)
County control variables		
Bank density	1.001	1.002
	(0.003)	(0.001)
Interactions		
Rural × 2009	0.589***	0.729**
	(0.072)	(0.086)
Rural × 2011	1.235	1.342**
	(0.129)	(0.142)
Rural × 2019	0.362***	0.449
	(0.064)	(0.061)
Year Dummies	Yes	Yes
Observations	5320	5433

LR $chi^2(18)$	882.93	445.03
Prob > chi^2	0.0000	0.0000
Pseudo R^2	0.5391	0.5646
Log likelihood	-2731.19	-3223.76

Notes: ***, **, *, stand for significance at 1%, 5% and 10% respectively.

Table 2.11 focuses on the relationship between the household and county control variables of unbanked rural and urban households and the credit use and transaction use of AFSs. Columns 1 and 2 show the relationships of AFS credit use and AFS transaction use with the predictor variables in unbanked households. Household respondent's age, number of persons in the household, having a college degree, employed, U.S. born or foreign born citizen, Black, Hispanic and Asian households compared to the reference category white, married, homeowner and previously banked characteristics show significant associations for AFS credit uses of unbanked rural households. Also, age, number of persons in the household, college degree, employed, U.S. born or foreign born citizen, Asian, married, home owner and previously banked variables show significant effects with AFS transaction use of unbanked households.

The binomial regression results indicate that the proportional odds ratio for age of unbanked households explains that 0.981 times lower to use AFS credit services and decrease with the age. Likewise, unbanked urban households use of AFS transaction services 0.989 times lower, given the other variables are held constant. The proportional odds ratio for number of persons in the household, 1.112 times greater for AFS credit use and 1.044 times greater for AFS transaction use of unbanked households. Also, having a college degree the odds on using AFS credit services is 0.567 times lower and using AFS transaction services is 0.509 lower for unbanked households. The proportional odds ratio being employed 1.140 times greater for using credit services and 1.381 times greater for transaction uses of unbanked households. The odds ratios for US born or foreign born citizens show positive significant effect and 3.157 times greater using credit services and 1.191 times greater for transaction services. Also, Black, Hispanic and Asian households tend to use more credit services and Asian households use more transaction services compared to White households. Married unbanked households 1.255 times more use AFS credit services and 1.215 times more use transaction services. The odds ratios for being a homeowner and previously banked show positive significant effects on AFS credit uses as 0.666 and 2.647 are the greater odds and for AFS transaction uses 0.740 and 2.148 are the greater odds of unbanked households. The effect on using AFS credit services by the unbanked rural household is 0.589 times lower and using AFS transaction services is 0.729 times lower compared that to unbanked urban household in 2009. Also, the effect on using AFS transaction services by unbanked rural household is 1.342 times greater in 2011. In 2019, unbanked rural households 0.362 times lower using credit services and AFS transaction services use is not significant in 2019. Both AFS credit and transaction services less likely used in 2009, but in the short term effect only in use of AFSs for transactions more likely used by rural relative to urban households, while the long term effect in use of AFS credit for 2019 is the opposite, rural households less likely to use AFS for credit in long term relative to urban households but no long term transaction AFSs used. The effects on year dummies are present in both AFS credit and transaction uses.

Conclusion

This paper set out to determine the impact analysis of financial inclusion of unbanked rural households after the Dodd-Frank act in 2010 using the Quantile treatment effect (QTE) with the Changes-in-changes (CIC) model considering 2009, 2011 and 2019 household data sets from a National household survey conducted by FDIC. The discrete CIC model developed using the common Difference in Differences (DID) Assumption and estimated the quantile treatment effects for the treated group unbanked rural households to identify the Average treatment effects of treated group by percentile. Unbanked household demographic and county variables are used to examine the impact on likelihood of opening bank account and using alternative financial credit and transaction services. Results indicate that the Dodd-Frank act is associated with higher likelihood of opening a bank account by the unbanked rural households' in both short and long term with an expected smaller long term magnitude. Relative to urban unbanked households, rural unbanked households are more likely to use AFS for transaction purposes in 2009 and even more in 2011, indicating a strong potential effect related to substituting transacting via banks with transacting via AFS. However, relative to urban households, rural households are more likely to have used more AFS for credit purposes, in the long run – in 2019, which may be related to the resulting closures of banking infrastructure in rural areas reducing bank branches density 32 to 24 per 100,000 people from 2009 to 2019.

As policy implications, Dodd-Frank act has imposed changes in banking and AFS sector between 2010 and 2019 and it is associated with to unbanked rural population more likely to shift towards using formal banking services. Although the demand is there, the closing of rural and community bank branches that happened during the last decade with likely negatively affect for these populations even the institutions provide cheaper services. To address this issue effectively, remaining rural and community financial institutions need practical processes and cost structures such as introduction of mobile banking technology or promoting online banking in rural communities. With better tools, community banks and credit unions can operate more efficiently, and provide better services for rural community who need banking services.

Growth of using mobile devices among communities for the past decade can be used as a digital channel to promote banking services, especially if offered without fees. The ability to open accounts online is a comparatively cost-effective way to reach unbanked rural community and it can also change the cost structure of financial institutions and improve efficiency, allowing for example to waive fees, or offer competitive loan interest rates to customers, as well as offer education to improve financial literacy and engage in marketing in unbanked rural communities to advance financial inclusion. With online account opening, rural and community bank institutions can make a real difference by creating opportunities for more unbanked people to enter formal banking system in near future.

References

Aportela, F. (1999). Effects of Financial Access on Savings by Low-Income People. Mimeo, Banco de México.

Ashraf, Nava, Dean Karlan, and Wesley Yin. (2006). "Deposit Collectors," *Advances in Economic Analysis & Policy* 6: 1–22.

Athey S, Imbens GW (2006). Identification and inference in nonlinear difference in difference. Econometrica 74(2):431–497

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. (2010). "The Miracle of Microfinance? Evidence from a Randomized Evaluation." MIT Bureau for Research and Economic Analysis of Development Working Paper no. 278. Massachusetts Institute of Technology.

Barr, Michael S. (2004). "Banking the Poor." University of Michigan Law School, *Law & Economics Working Papers Archive: 2003–2009*. Available at https:// repository.law.umich.edu /law_econ_archive/art49/

Beatty, T. K., & Shimshack, J. P. (2011). School buses, diesel emissions, and respiratory health. Journal of health economics, 30(5), 987-999.

Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine. (2007). "Finance, inequality and the poor." *Journal of Economic Growth* 12(1): 27-49.

Berry, Christopher. (2004). "To Bank or Not to Bank? A Survey of Low-Income Households." Joint Center for Housing Studies Working Paper Series, BABC 04-3. February 2004.

Board of governors of the federal reserve system. (2019). Perspectives from Main Street: Bank Branch Access in Rural Communities. Retrieved from: https://www.federalreserve.gov/ publications/ files/bank-branch-access-in-rural-communities.pdf

Brune, Lasse, Xavier Giné, Jessica Goldberg, and Dean Yang. (2011). "Commitments to save: A field experiment in rural Malawi." *World Bank Policy Research Working Paper Series*, Vol (2011).

Burgess, R. and R. Pande (2005). Can Rural Banks Reduce Poverty? Evidence from the Indian social banking experiment. *American Economic Review* 95(3), 780–795.

Campbell, D., F.A. Martinez-Jerez and P. Tufano (2012). Bouncing Out of the Banking System: An Empirical Analysis of Involuntary Bank Account Closures. *Journal of Banking and Finance* 36, 1224-1235.

Caskey, John P. (2002). "Bringing Unbanked Households into the Banking System." *Capital Xchange*, January. Available at https://www.brookings.edu/wp-content/uploads/2016/06/ caskey.pdf

Célerier, C., & Matray, A. (2019). Bank-branch supply, financial inclusion, and wealth accumulation. *The Review of Financial Studies*, *32*(12), 4767-4809.

Chernozhukov, V., FernandezVal, I., Hahn, J., Newey, W., (2013b). Average and quantile effects in nonseparable panel models. Econometrica 81 (2), 535–580.

Chernozhukov, V., Hansen, C., (2006). Instrumental quantile regression inference for structural and treatment effect models. J. Econometrics 132 (2), 491–525.

Council of Economic Advisors Issue Brief (2016). Financial Inclusion in the United States. Retrieved from: https://obamawhitehouse.archives.gov/sites/default/ files/docs/ 20160610_ financial_inclusion_cea_issue_brief.pdf

Covington, M. and J. Courtney (2014). Banking in Rural America, Insight from a Community Development Financial Institution (CDFI).

D'Haultfoeuille, X., Hoderlein, S., Sasaki, Y., (2015). Nonlinear Difference-in-Differences in Repeated Cross Sections with Continuous Treatments, Working Paper.

De Alwis, D. (2020). Distributional impacts of disaster recovery: Sri Lankan households a decade after the 2004 Indian Ocean tsunami. Economics of Disasters and Climate Change, 4(1), 195-222.

Demirgüç-Kunt, A. and L. Klapper (2013). Measuring Financial Inclusion: Explaining Variation in Use of Financial Services across and within Countries. Brookings Papers on Economic Activity Spring 2013, 279-340

Dunham, I. M. (2019). Landscapes of financial exclusion: Alternative financial service providers and the dual financial service delivery system. *Business and Society Review*, *124*(3), 365-383.

Dupas, P. and J. Robinson. (2013). "Why Don't the Poor Save More? Evidence from Health Savings Experiments," *American Economic Review*, 103: 1138–71.

Federal Deposit Insurance Corporation. (2017). National Survey of Unbanked and Underbanked Households. https://www.economicinclusion.gov/

Fernandes, Daniel, John G. Lynch Jr, and Richard G. Netemeyer. (2014). "Financial literacy, financial education, and downstream financial behaviors." *Management Science* 60(8): 1861-1883.

Geler, B. (2020). What the Dodd-Frank Act Did (and How It's Changed). Retrieved from: https://smartasset.com/ investing/dodd-frank-act

Hogarth, Jeanne M., Christoslav E. Anguelov, and Jinhook Lee. (2005). "Who Has a Bank Account? Exploring Changes Over Time, 1989–2001." *Journal of Family and Economic Issues*, vol. 26, no. 1, pp. 7–30. Available at https://doi. Org/10.1007/s10834-004-1410-6

Karlan, Dean and Jonathan Zinman. (2010). "Expanding Credit Access: Using Randomized Supply Decisions to Estimate the Impacts." *Review of Financial Studies* 23(1): 433–64.

Karlan, Dean, Aishwarya Lakshmi Ratan, and Jonathan Zinman. (2014). "Savings by and for the Poor: A Research Review and Agenda." *Review of Income and Wealth* 60(1): 36-78.

Koenker, R., Bassett, G., (1978). Regression quantiles. Econometrica 33–50.

Le, H. (2017). The Effects of the Dodd-Frank Act on Community Banks.

Melly B, Santangelo G (2015). The changes-in-changes model with covariates (Working Paper)

Osili, U.O. and A. Paulson (2008). What Can We Learn about Financial Access from U.S. Immigrants? The Role of Country of Origin Institutions and Immigrant Beliefs. *World Bank Economic Review* 22(3), 431–455.

Rhine, S.L.W. and W.H. Greene (2006). The Determinants of Being Unbanked for U.S. Immigrants. *Journal of Consumer Affairs* 40(1), 21–40.

Rhine, S.L.W. and W.H. Greene (2013). Factors That Contribute to Becoming Unbanked. *Journal of Consumer Affairs*, 47(1), 27–45.

Ruiz Ortega, Claudia. (2013). "From pawn shops to banks: The impact of formal credit on informal households." *World Bank policy research working paper 6634*.

Schorgl, M. (2018). *The Dodd-Frank Act and its Impact on Community Bank Mortgage Lending* (Doctoral dissertation, University of Mississippi).

Wheelock, D. (2012). Too big to fail: The pros and cons of breaking up big banks. The Regional Economist. Federal Reserve Bank of St. Louis. Available at https://www.stlouisfed.org/publications/re/articles/?id=2283

Appendix A

Table A.1. Average Treatment effect (ATE) for the outcomes of unbanked rural households without covariates

Outcome variable	2011			2019		
	Average Effect	S.E.	Obs.	Average Effect	S.E.	Obs.
Like to open a Bank account	-0.80	0.066	6190	-0.18	0.095	4619
AFS Credit services use	-0.60	0.027	6190	-0.14	0.026	4619
AFS Transaction services use	-0.26	0.022	6190	-0.17	0.022	4619

Note: ***, **, *, stand for significance at 1%, 5% and 10% respectively.

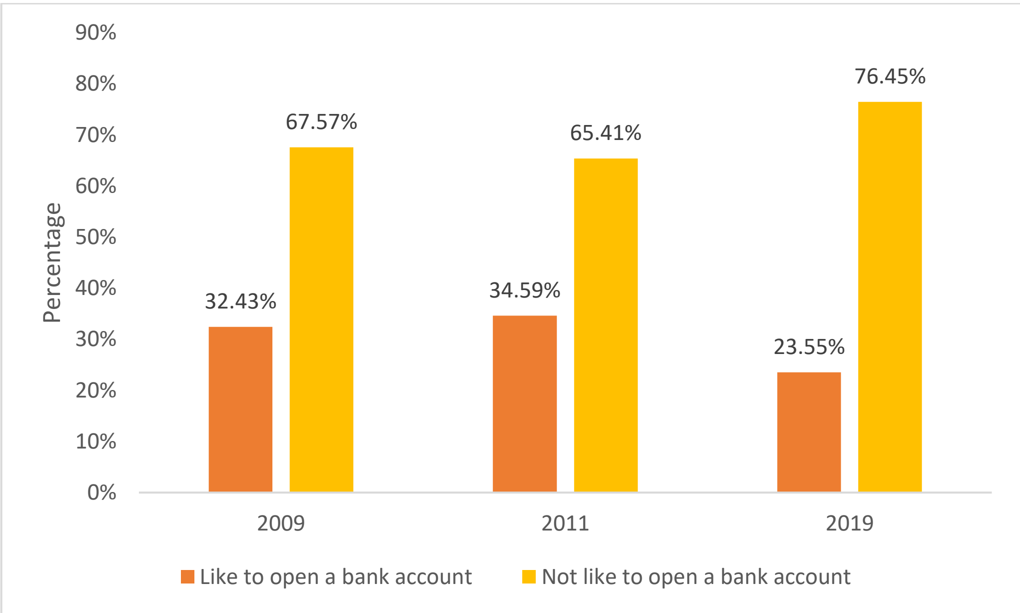

Figure A.1 Likelihood of opening a bank account of Unbanked households from 2009 – 2019

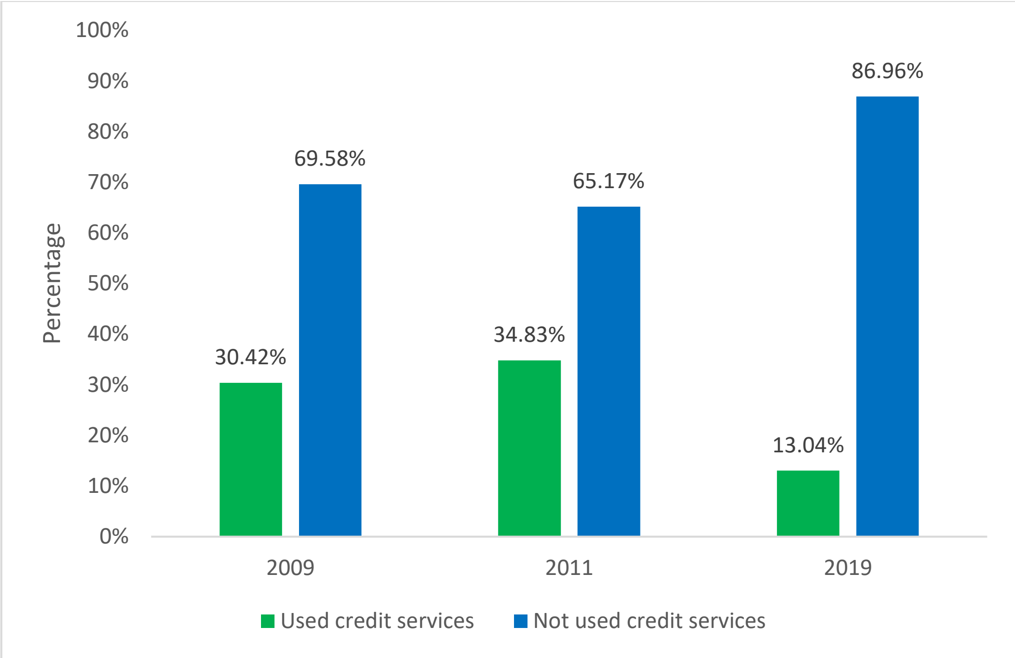

Figure A.2 Alternative Financial Services (AFS) Credit use of Unbanked households from 2009 – 2019

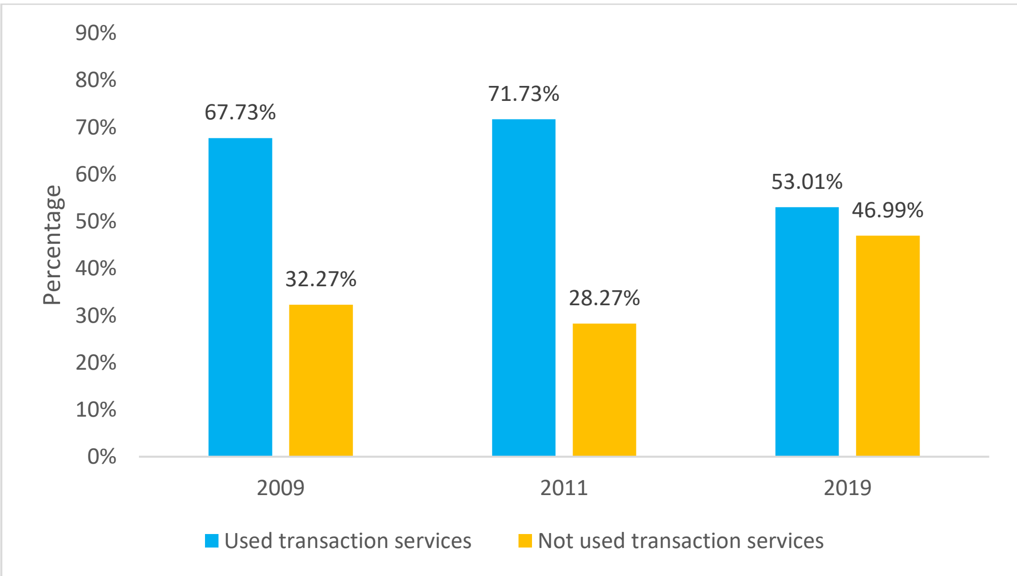

Figure A.3 Alternative Financial Services (AFS) Transaction use of Unbanked households from 2009 – 2019

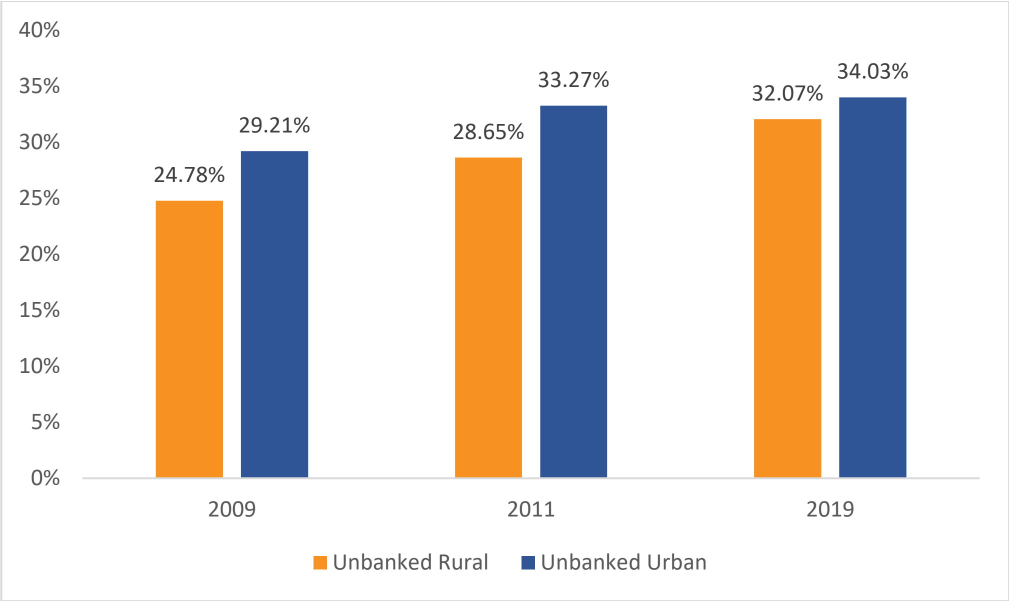

Figure A.4 Unbanked Urban and Rural Households' percentage like to open a bank account from 2009 – 2019

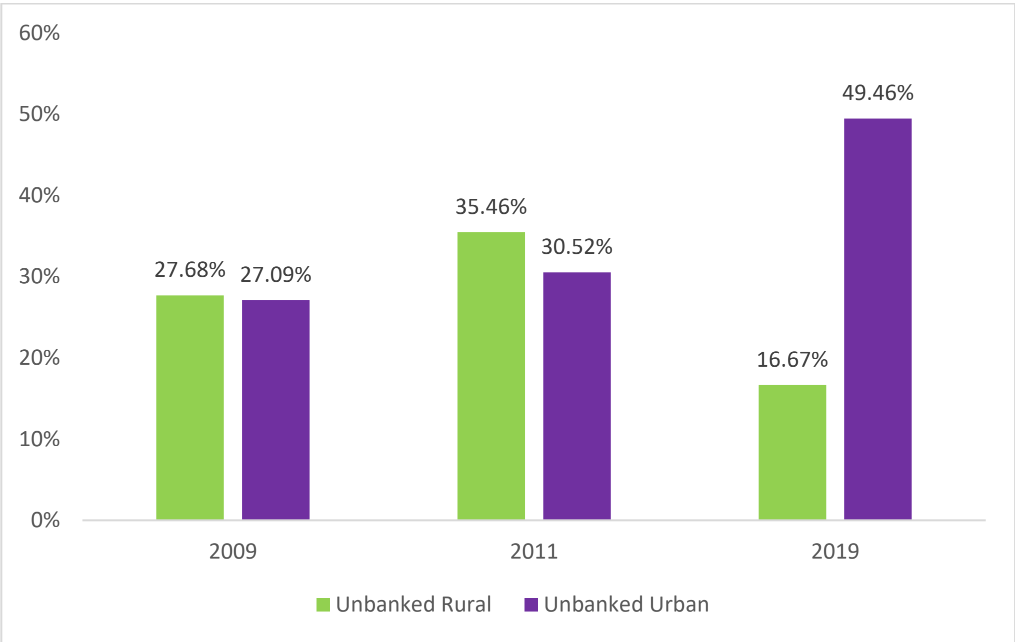

Figure A.5 Unbanked Urban and Rural Households' use of Alternative Financial Services (AFS) Credit services from 2009 – 2019

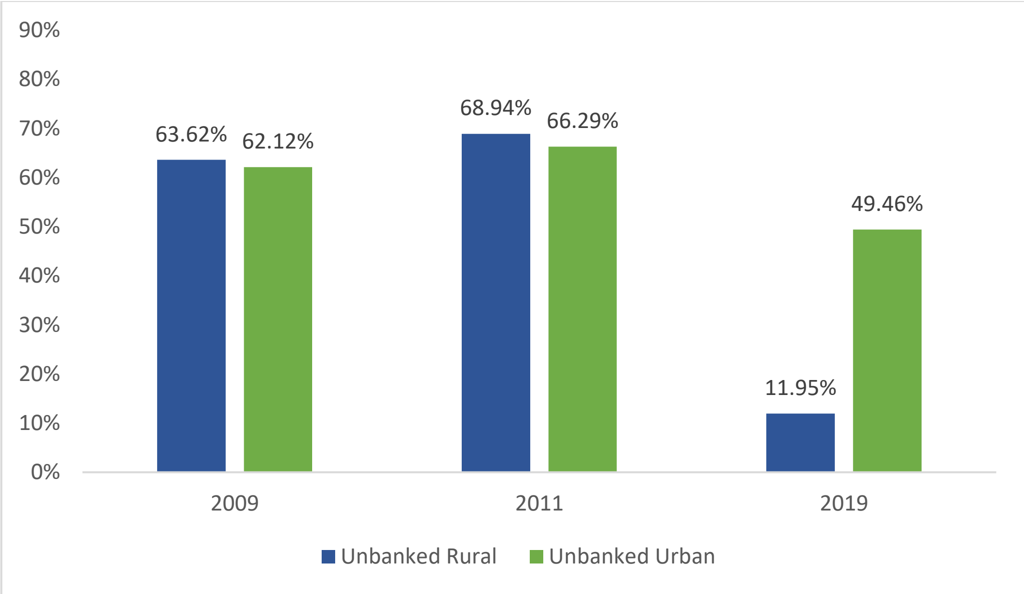

Figure A.6 Unbanked Urban and Rural Households' use of Alternative Financial Services (AFS) Transaction services from 2009 – 2019

3. CHAPTER 3

Prepaid cards use and its impact on Financial inclusion and Alternative Financial services among Unbanked households

Introduction

With the development of financial services in developed economies, access to financial services for households has become an important concern on a global scale. Prepaid cards have become one of the fastest growing payment method that helps consumers to fulfil their financial needs. The prepaid card industry has grown rapidly since the financial crisis in 2008 because these debit cards fill in the gap of retail operations in the traditional banking sector for the past decade (Zywicki, 2013). Prepaid cards are a financial service option available for unbanked households and they are particularly important as an electronic payment method alternative to traditional bank transactions and credit card payment.

Based on the latest data extracted from the 2019 FDIC National Survey using 33,000 households, 5.4 percent households were unbanked in the United States (US). This percentage indicates that household or any person in the household had no checking or savings account and this fact represents approximately 7.1 million US households. Unbanked rate of 5.4 percent is the lowest since 2009 and that was 7.6 percent, around 9 million. Between 2009 and 2019, unbanked rate 7.6 in 2009 and peaked at 8.2 percent in 2011, then gradually dropped by 2.8 percentage points until 2019, corresponding to an increase of nearly 3.7 million banked households as shown in figure 3.1. Between 2011 and 2019, this decline is about two-thirds of the unbanked rate, and related with improvements of socioeconomic status such as higher annual income, lower monthly income uncertainty, higher rates of home ownership status, low unemployment and educational status of US households between 2009 and 2019 period (FDIC, 2019).

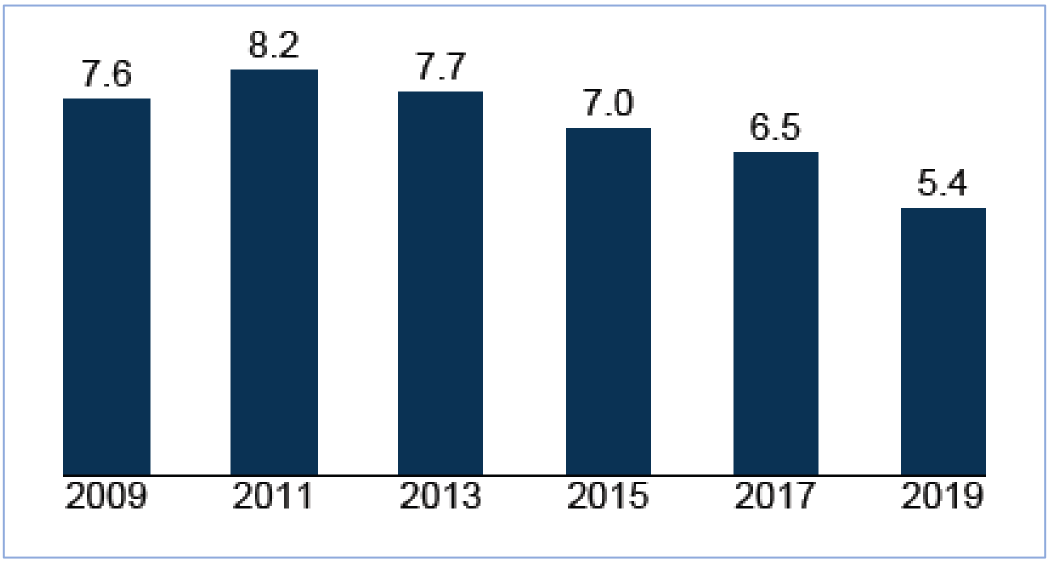

Figure 3.1. Unbanked Household percentage for 2009 – 2019 FDIC survey period
Source: FDIC Household Survey Report, 2019.

The main reason that do not have a bank account for the half of unbanked households was not enough money to meet the minimum balance requirements to keep up an account (FDIC, 2019). Although the unbanked households use in person alternative transaction and credit services such as check cashing, money orders and loans, they are unable to participate in online and mobile transactions, which have become common and trending in payment systems (Toh, 2021). Lack of a bank account to a consumer is a crucial barrier for electronic payments and unbanked households use prepaid cards as the common method for electronic payments. Consumers use General purpose reloadable prepaid cards as a common payment type for financial transactions such as making purchases, paying bills, withdrawing cash at ATMs, depositing checks, and receiving direct deposits etc. (FDIC, 2019). In 2019, percentage of unbanked households use a prepaid card is 27.7 and banked households and prepaid card use was more common for unbanked households than for banked households, consistent with survey results for the previous 10 years. The decline rate of unbanked households, together with the growth in the prepaid card industry shows that people may use prepaid cards to reconnect with the retail banking sector and other essential FinTech services.

Most of the unbanked households, regardless of whether they use or not use prepaid cards consider using other non-traditional bank options such as Alternative Financial Services (AFS) to take care of their financial credit and transaction requirements. The use of AFS include credit (payday loans, auto title loans etc.), transactions (check cashing, money orders etc.), investment and insurance from informal financial institutions and are considered an indicator of financial exclusion (Lamb, 2016). FDIC surveys show that the proportion of unbanked households that used both AFS credit and transaction services has declined and AFS use is much higher among unbanked households than banked households in recent years (FDIC survey, 2017), half of unbanked adults used some form of AFS during 2019 period and 88 percent used AFS transaction services such as purchasing a money order or cashing a check. Twenty-nine percent used AFS for credit namely payday loans or paycheck advances, auto title or pawn shop loans, and tax refund advances.

Even though unbanked households use prepaid cards as a substitute for traditional financial services, it is important to understand whether this trend of using prepaid cards is helping the unbanked households to access banking services in the future. Earlier literature has broadly studied the determinants of being unbanked. However, there are limited studies that explore how prepaid debit card use may influence future financial access or impact on AFS use of the unbanked households. This paper fills in the gap of the literature in particular, it investigates the prepaid debit card use of unbanked households impact on financial inclusion and alternative financial services. I employ the Propensity Score Matching (PSM) to identify the causal impact of the how the use of prepaid card affects financial inclusion as well as in understanding how the use of prepaid cards affects the use of AFS for credit and transaction purposes.

This study uses household data from the FDIC National Surveys of Unbanked and Underbanked Households supported to the Current Population Survey (CPS) of US Census

Bureau and consists of annual surveys for the period 2009-2019. The data contains demographic, economic, and financial data. The study focuses on the unbanked households and has the following objectives: (1) identify how plans to open a bank account is affected by the use of prepaid debit cards, (2) identify how use of credit AFS is affected by use of prepaid cards and (3) how use of AFS for payment and transaction purposes is affected by the use of prepaid debit card of unbanked households. This paper is organized as following sections. Next section provides a literature review and follows the data description and methodology related to the econometric approach. The fourth section discusses the results and discussions and the last section provide the conclusion and plans for further implications of the study.

Literature Review

Financial inclusion is an important and popular topic in the economic literature for the last decade. Financial inclusion and financial access are closely related ideas. The goal of financial inclusion is to provide access to financial services to all unbanked households who can use these services to improve their living standards (Kim, 2016). There are many definitions to financial inclusion because it involves complex societal behavior. Typically, financial inclusion is defined as a measure of having access to financial products and services to meet the needs of individuals and businesses such as credit, transactions, savings, payments, and insurance (World Bank,2018). According to the Federal Deposit Insurance Corporation (FDIC), financial inclusion is defined as efforts by the public and private sectors aimed at bringing unbanked and underbanked people who access only alternative financial services (AFS) into the formal finance sector .

Over the last two decades, financial inclusion has become an important discussion of development economics at the same level as education, healthcare, property rights and infrastructure as an important factor that can contribute to the economic growth and reduce poverty (Karp and Nash-Stacey, 2015). Several studies have shown a link between financial services' growth and financial inclusion, and it facilitates to deal with more financial expansion in the United States. Demirgüç-Kunt and Klapper (2013) find significant positive correlation between adults who have accounts with formal financial institution, and the level of the country's domestic credit. They also find that that financial inclusion more broadly correlated with economic development.

There are a number of barriers to financial inclusion regardless of the benefits. Ashraf, Karlan, and Yin (2006) showed that there is a potential barrier with upfront cost and other fees associated with opening a bank account, minimum balance requirement to keep up the account, requirement of necessary documents and proof of identity, cost related with traveling to a bank branch, and opportunity costs linked with opening a bank account. Other potential barriers such as mistrust of banks and lack of financial capability related to financial inclusion is not entirely clear (Karlan, Ratan and Zinman 2014; Fernandes, Lynch and Netemeyer 2014).

FDIC survey reported from the data of unbanked households in 2019 that 56.2 percent of the unbanked were not at all interested in having a bank account, while 24.8 percent were very or somewhat interested to have a bank account. These estimates are similar to the data from the 2017 survey to some extent, showing that 58.7 percent were not at all likely, 16.3 percent were not very likely, 25.1 percent were somewhat and very likely of the unbanked to open a bank account in the next 12 months. The questions in the 2019 survey are not identical to those in 2013 to 2017 . In 2019, the question asked if the unbanked households were interested in having a bank account, while in the other surveys the question was how likely was it that the household will open a bank account in the next 12 months (Federal Deposit Insurance Corporation, 2019). According to the 2019 report of economic wellbeing of US households by the Board of Governors of the Federal system, the rate of unbanked adults was 6 percent, which is close to the value of FDIC surveys.

The findings in the FDIC's National Survey are similar to what other studies have found of Unbanked Households in previous years. Unbanked consumers are more likely to be younger, less formal education, are of a racial or ethnic minority, more likely to be disabled, and lower-income or incomes with considerable monthly variation unlike the incomes of the general US population (Congressional research service, 2019). Sherraden (2010) discovers that young, less educated, the minorities, unmarried, and those less wealthy do not intend to have checking or savings accounts. The unbanked people themselves are more likely to be younger, less educated, low income, immigrant, and female head of household (Beard, 2010). Vermilyea and Wilcox (2002) found that income, age, spending habits, education, race/ethnicity, home ownership, and employment are significantly associated with the participation in the formal banking system. In New York, thousands of low income families choose to stay unbanked because of excessive banking fees, and they not have the educational awareness to choose banking services (NY Neighborhood Financial Services Study, 2008). Also, factors such as educational background, income, English language proficiency, legal status, and longevity of residence were found to impact the decision to remain unbanked or underbanked in immigrant communities (Meghan, 2008). Cole and Greene (2016) used the Survey of Consumer Payment Choice (SCPC) and examined the relationship between consumers' banking status, and their sociodemographic characteristics using data, and found that the lowest income is correlated with being unbanked. This survey is conducted by the Federal Reserve Bank of Boston.

Unbanked individuals who live in a household without a checking or a savings account at a bank or FDIC certified institution typically use alternative financial services (AFSs) to meet their banking needs. These AFSs offer check cashing, money orders, payday, auto title, pawn shop, and refund anticipation loans, and rent-to-own services same as formal banking services (Bradley et al., 2009). AFSs are used the most by unbanked households and being unbanked highly correlated with high-cost borrowing, including pawn shops and rent-to-own shops, payday and refund anticipation loans, with financial vulnerability (Lusardi and Scheresberg, 2013). Smith et al. (2008) analyzed the locations of AFS providers and found that they are located in places that lack access to traditional banks. Birkenmaier and Fu (2016) studied heterogeneity in the use of AFS and found that being unbanked was significantly associated with the use of AFS, both credit and transactional services. Recent studies have found that unbanked consumers use AFS services than banked population. In 2019, 42.3 percent money orders, 31.9 percent check cashing, and 14.4 percent used bill payment services combining more than half of the unbanked households (56.1 percent) used at least one of these three types of services. Additionally, 9.4 percent international remittances, and 8.8 percent used person to person payment services indicating that Unbanked households use AFS transaction services more than AFS credit services (Federal Deposit Insurance Corporation, 2019).

Unbanked households are more likely to use prepaid cards compared with banked households. Also, changes in prepaid card usage in earlier years were similar to the changes among unbanked in 2019. Prepaid card usage was higher among younger, less educated, Black, lower or unstable income, working-age disabled households. Hayashi (2016) compared the costs of general-purpose reloadable (GPR) prepaid cards, bank checking accounts, and AFS data and found that the cost of GPR prepaid cards strongly dependent on overdraft behavior of unbanked people and they find that these cards are a cheaper option compared to bank accounts.

Anong and Routh, (2021) found that there is no empirical association between prepaid debit card use and banking intention using an unbanked sample from 2012 to 2017. Also, they found that current prepaid card users were equally likely to be recently banked or to be long term unbanked but less likely to be long term banked focusing on another sample. Also, they found factors such as use of other AFS transactions and credit services, more recent relationship with banks, having a smartphone and trust of banks increase the banking intention. Also, a recent study by Toh (2021) used a data sample of unbanked households from the 2019 FDIC survey, to examine whether the use of prepaid cards improves digital payment inclusion and found that prepaid card adoption rate has been low among unbanked households. Both being unbanked and using a prepaid card was due to same reasons; privacy concerns, lack of trust in the formal banking institutions, liquidity constraints and personal identification issues (Toh, 2021). The descriptive statistics of the same survey in 2019 show that unbanked households who used prepaid cards were more likely to have had a bank account compared to unbanked households who did not use prepaid cards. In 2019, 65.3 percent of the prepaid card users had previously been banked compared to 45.1 percent of unbanked who did not use prepaid cards. In addition, unbanked households that used prepaid cards were more interested in having a bank account than who did not use prepaid cards. Also, the survey shows 30.6 percent who used prepaid cards were very or somewhat interested in having a bank account, compared with 22.8 percent who did not use prepaid cards of unbanked households.

This study fills the gap in the financial inclusion literature by identifying the relationship between the likelihood of opening a bank account, changes in using AFS credit and transaction services and prepaid debit card use, exploring the FDIC Household Surveys of the Unbanked and Underbanked for the period 2009 to 2019 . The surveys contain comprehensive data of households' financial services and socio-demographic characteristics, and their reasons for being unbanked for both unbanked prepaid debit card users and non-users. This idea allows to study how use of prepaid card affects plans to open a bank account, as well as the use of alternative financial services such as credits and transactions for nationally representative households.

Data and Methodology

This paper relies on the survey data from the FDIC National household survey from the years 2009 to 2019. This survey data contains information such as plans for opening a bank account in the future, use of Alternative Financial Services (AFS) of credits and transactions, households' demographic status and socioeconomic status. Responses of the possibility of opening a bank account are from the question "how likely the respondent or someone in the household is to open a bank account within the next 12 months." Also, to measure of using AFS, utilizes the questions if the respondent or person in the household has ever gone to a non-formal location for credit and transaction services or used credit and transaction services for the past 12 months. Prepaid debit card using information comes from a question about use of prepaid debit cards that have logos such as VISA, MasterCard, Discover or American Express that are not linked to a checking or savings account. These cards can be used to make purchases and pay bills anywhere that credit cards are accepted or withdraw the cash from an ATM (Automated Teller Machine) or to add money onto the card. There are 14756 observations of unbanked households between 2009 to 2019, but 522 households were dropped due to the missing key data on the matching variable prepaid card use information, resulting into 14234 observations

The propensity score matching approach (PSM) applied in this study to measure the effect of prepaid card use on financial inclusion, measured as the intention/likelihood to open a bank account. Financial inclusion effect of the prepaid card using households was evaluated by controlling for various aspects of household characteristics and county characteristics. The outcome variables, measure financial inclusion, is a binary indicator taking the value of 1 if a household plans to open a bank account and 0 otherwise.

The specific explanatory variables included in the model are include the following demographic and socio-economic characteristics; alternative financial services (AFS) credit use, AFS transaction services use, household respondent's age, number of persons in the household, Spanish only language spoken, family income less than $15,000, high school diploma and college degree in education, employment status, citizen status, Black and White in race, Married and female headed family as household types, home ownership and previously banked status of households. Household respondent's age and number of persons in the household included as continuous control variables and other all variables contained as dummy explanatory variables. All analyses were conducted using the Stata 16.0 version software.

Also, answers to the questions related to the use of AFS credit services such as payday loan, pawnshop loan, rent-to-own service, and refund anticipation loans and using AFS transaction services such as check cashing and money order and responses to this question were used to create a binary indicator equaling 1 if a household indicated that they used AFS credit services or transaction services and 0 otherwise. Following previous studies, the present study identified 17 covariates with the dependent dummy variable - Like to open a bank account for the use of prepaid card as a measure for financial inclusion and later in the study used two secondary outcomes of interest - AFS services for credit and for transaction purposes respectfully for additional explanation. Table 3.1 provides a description of the covariates used in the study.

Table 3.1 The definition of variables including household characteristics and outcomes of Unbanked prepaid card using and not using households

Variable	Variable description
Like to open a bank account	1 = Very like to open a bank account
	0 = Not at all like to open a bank account
AFS Credit use	1 = Household use AFS credit services
	0 = Household do not use AFS credit services
AFS Transaction use	1 = Household use AFS transaction services

	0 = Household do not use AFS transaction services
Age	Household respondent's age
Number of persons in the household	Number of people who live in the household
Spanish only language spoken	Spanish is the only language spoken in the household
Rural	1 = Household live in the rural area
Family income less than $15k	1 = Family income of the household is less than $15,000
Highschool diploma	1 = Education level of the household is high school diploma
College degree	1 = Education level of the household is a college degree
Employed	1 = Household is employed
US born / foreign born citizen	1 = Household is a US born or foreign born citizen in US
Black	1 = Race of the respondent known as Black
White	1 = Race of the respondent known as White
Married	1 = Marital status of the household is married
Female headed family	1 = Household type identified as a female head
Homeowner	1 = Respondent is identified as a home owner
Previously banked	1 = Household had used bank services or an account earlier

Rosenbaum and Rubin (1983) established propensity score matching (PSM) method, used later by others who adapted and advanced its applications (Dehejia and Wahba, 1999;2002; Becker and Ichino, 2002 and Caliendo and Kopeinig, 2005). These papers defined propensity score as a conditional probability of treatment given by pre-treated characteristics of the subject. They argued that since the choice of subjects to be treated or in the control groups in a given treatment may not be random, then the estimation of the treated effect may be biased by the existence of confounding factors. Therefore, they proposed PSM method for the comparison of treated and control subjects who are as similar and correct for the estimation of effects of the treatment controlling for the existence of these confounding factors when the bias is reduced. The propensity score is a method that summarize pre-treated characteristics and a conditional probability estimator that use as any discrete choice model such as logit or probit while they generate similar results (Caliendo and Kopeinig, 2005).

PSM has been used as an alternative method to estimate causal treatment effects of balanced treatment groups to assess the treatment effect on the outcome with reduced bias. As described in equation (1), in the PSM, Y_{i0} and Y_{i1} are the outcome variables describing household's desire to open a bank account and AFS usage patterns for unit i conditional on the treatment and control respectively. T_i denote as an indicator variable equal to 1 if individual i is a treated individual (prepaid card using unbanked household) and 0 if unbanked household is not using a prepaid card. The treatment effect for individual i (ΔY_i), measures the difference between the relevant outcome indicator with the treatment (prepaid card use) and without the treatment. It is shown as:

$$\Delta Y_i = E(Y_{i1}/T_i = 1) - E(Y_{i0}/T_i = 1) \quad (1)$$

When the value of post-treatment outcome for individual is observed, value of the pre-treatment is not observed. It is impossible to simultaneously observe an individual in two different states when the components $E(Y_{i1}/T_i = 1)$ and $E(Y_{i0}/T_i = 0)$ are observable outcomes, while $E(Y_{i1}/T_i = 0)$ and $E(Y_{i0}/T_i = 1)$ are non-observable outcomes. PSM method is based on the conditional independence assumption (CIA), which outcome in the untreated state is independent of treatment participation conditional on a vector of observable characteristics (Rosenbaum and Rubin 1983). This assumption is equivalent to the absence of selection bias based on unobservable heterogeneity (Heckman and Robb 1985).

$$(Y_{i0}, Y_{i1}) \perp T_i/X_i \quad (2)$$

Equation (2) express that, given X_i, the outcomes of nontreated units can be used to approximate the counterfactual outcome of treated units in the absence of treatment.

$$E(Y_{i0}/T_i = 1, X_i) = E(Y_{i0}/T_i = 0, X_i) \quad (3)$$

The propensity score $P(X_i)$ is the conditional probability of receiving a treatment (Prepaid card use) given pre-treatment characteristics or household characteristics.

$$P(X_i) = Prob(T_i = 1|X_i) = E(T_i|X_i) \quad (4)$$

where D = (0, 1) is the binary variable on whether an unbanked household used prepaid cards (1) or not (0) and X is the multidimensional vector of pre-treatment or time-invariant or relatively stable characteristics in the framework. Rosenbaum and Rubin (1983) established a balancing hypothesis and unconfounded conditions in order to estimate Average Treatment on the Treated (ATT) effect based on the propensity score as follows:

$$ATT = \{Y_{i1} - Y_{i0} \mid T_i = 1\}$$

$$ATT = E\big\{E\{Y_{i1} - Y_{i0} \mid T_i = 1, P(X_i)\}\big\}$$

$$ATT = E\{E\{Y_{i1} \mid T_i = 1, P(X_i)\} - E\{Y_{i0} \mid T_i = 0, P(X_i)\}| \, T_i = 1\} \quad (5)$$

where i denotes the *i-th* household, Y_{i1} is the potential outcome in the two counterfactual situations with and without the effect of prepaid card use. However, various matching methods use to estimate the ATT using Equation (5) because the estimation of the propensity score is not enough and the probability of observing two units is zero with exactly the same value of the propensity score since $P(X_i)$ is a continuous variable (Oh et al. 2009).

The nearest neighbor matching (NNM) method also called as greedy matching used in this study. NNM use to match each treated individual to the nearest untreated individual, that is individuals with closest propensity scores are matched. A distance measure condition is used to define which control unit is closest to each treated unit. The default method of distance is the propensity score difference, which is the difference between the propensity scores of each treated and control unit used in this study (Stuart, 2010). NNM technique is used to compare treatment groups with respect to control group characteristics. The traditional NNM randomly selects a treatment unit and a control unit for each treatment unit with the smallest distance that can be run with or without replacement. With replacement, a control unit is return back after a match and can be matched later to another treatment unit. Without replacement, a control unit is taken out once it is matched (Morgan and Harding, 2006). This propensity scores are estimated based on logit or probit model, and thus estimating the average treatment effect of the outcome difference between the treatment (prepaid card users) and the control (non-prepaid card users) groups using nearest-neighbor matching method.

Results and Discussion

Table 3.2 presents the mean summary statistics for the household variables of prepaid card using and not using unbanked household groups from 2009, 2011, 2013, 2015, 2017 and 2019. Also, table shows that households' financial inclusion and demographic characteristics across total unbanked household sample before matching.

Table 3.2. Mean summary statistics of household demographic variables for Unbanked prepaid card using and not using households

Household Variables	Total Unbanked household sample (n=14,234) Mean (SD)	Unbanked Prepaid card users (Treated, n=3,351) Mean (SD)	Unbanked Prepaid card non users (Control, n=10,883) Mean (SD)
Like to open a bank account	0.43 (0.49)	0.39 (0.49)	0.43 (0.49)
AFS Credit use	0.26 (0.44)	0.44 (0.50)	0.20 (0.39)
AFS Transaction use	0.62 (0.48)	0.77 (0.42)	0.57 (0.49)
Age	44.86 (16.18)	41.05 (13.70)	45.96 (16.64)
Number of persons in the household	2.56 (1.69)	2.72 (1.66)	2.51 (1.70)
Spanish only language spoken	0.09 (0.28)	0.03 (0.17)	0.10 (0.31)
Rural	0.22(0.41)	0.23 (0.42)	0.22 (0.41)
Family income less than $15000	0.56 (0.49)	0.53 (0.49)	0.57 (0.49)

Highschool diploma	0.38 (0.48)	0.40 (0.49)	0.37 (0.48)
College degree	0.05 (0.21)	0.04 (0.20)	0.05 (0.22)
Employed	0.41 (0.49)	0.45 (0.50)	0.40 (0.49)
U.S. born / foreign born citizen	0.84 (0.37)	0.93 (0.25)	0.81 (0.39)
Black	0.31 (0.46)	0.33 (0.47)	0.30 (0.46)
White	0.38 (0.48)	0.47 (0.50)	0.35 (0.48)
Married	0.21 (0.41)	0.19 (0.40)	0.22 (0.41)
Female headed family	0.28 (0.45)	0.36 (0.48)	0.26 (0.44)
Homeowner	0.24 (0.43)	0.19 (0.39)	0.26 (0.44)
Previously banked	0.49 (0.50)	0.66 (0.47)	0.43 (0.49)
Observations	14,234		
2009	2,776		
2011	2,954		
2013	2,505		
2015	2,322		
2017	2,066		
2019	1,611		

According to financial inclusion indicators, 43% of unbanked households like to open a bank account, 26% use alternative financial credit services and 62% use alternative transaction services through the whole sample. The average age of household in the sample is 44.86 years and a household has on average three people (2.56). Also, 9% of households spoke only Spanish language, 22% lived in rural areas and 56% of household's family income is less than $15k. In terms of educational attainment, 38% report having high school education and 5% attended college/university level. Nearly 41% of the respondents have employed and 84% of households reported as US born and foreign born citizens. From the racial background, black and white selected as the main racial groups as other races did not show significant change in the sample. 31% of households are black and 38% reported as white from the races in the total sample. Also, 21% are married, 28% are female heads of household, 24% are home owners and 49% are previously banked in this sample. In previously banked status, unbanked prepaid card users 66% reported as previously banked compared to 43% of unbanked prepaid card non using households.

In Table B.1 (in the Appendix), report the logit regression estimates of explanatory variables used in the logit model that used for propensity score matching. After controlling for household effects, results show that the prepaid card use, AFS credit use, AFS transaction use, household's age, Spanish only spoken, rural, family income less than $15k, being a U.S. born and foreign born citizen, female headed family households more like to involve with the financial inclusion with the propensity to use prepaid cards. Also, the t-test for the differences between users and non-users of prepaid cards reveal significant differences in terms of prepaid card use, AFS credit and transaction use, age, Spanish only spoken, family income less than 15k, U.S. born, and foreign born citizen and female headed family status. Ali these preliminary investigations point to the need to use a matching method.

The matching process starts with the propensity scores estimation for the treatment variable. Logistic regression model was used to find the probability of opening bank account was regressed to number of covariates. The results of estimation of the propensity scores are reported in the appendix Table B.2. These results indicate that many of the household's indicator variables influence the probability of opening a bank account and thus the future access to formal financial services. In particular, using AFS credit and transaction services, number of persons in the household, family income less than $15k, employed, US born or foreign born citizen, Black, White, female headed family and previously banked have positive coefficients and age, Spanish only language spoken, and home owner have negative coefficients to fit the model to adapt with financial services.

Table 3.3 presents the average sociodemographic characteristics between unbanked prepaid card users and propensity score-matched unbanked prepaid card non-users. Each of the 2609 prepaid card users was matched to their nearest neighbor, and thus 2609 prepaid card non-users were included in the analysis. The table shows that prepaid card users and non-users were significantly similar across age, number of people in the household, Spanish only language spoken, family income less than $15k, US born and foreign born citizen, Black, White, female headed family, home owner and previously banked sociodemographic variables.

Table 3.3. Household characteristics of Propensity score matched Unbanked prepaid card using and not using households

Household Variables	**Unbanked Prepaid card users** (Treated, n=2609) Mean (SD)	**Unbanked Prepaid card non users** (Untreated, n=2609) Mean (SD)	**P-value**
Like to open a bank account	0.41 (0.49)	0.46 (0.50)	-
AFS Credit use	0.49 (0.50)	0.22 (0.41)	0.000
AFS Transaction use	0.79 (0.41)	0.60 (0.49)	0.000
Age	40.52 (13.54)	44.96 (16.10)	0.000
Number of persons in the household	2.77 (1.66)	2.56 (1.71)	0.000
Spanish only language spoken	0.03 (0.18)	0.11 (0.31)	0.017
Rural	0.24 (0.43)	0.23 (0.42)	0.406
Family income less than $15000	0.54 (0.50)	0.57 (0.49)	0.044
Highschool diploma	0.40 (0.49)	0.37 (0.480	0.670
College degree	0.04 (0.20)	0.04 (0.20)	0.156

Employed	0.44 (0.50)	0.40 (0.49)	0.189
U.S. born / foreign born citizen	0.93 (0.25)	0.81 (0.40)	0.000
Black	0.33 (0.47)	0.30 (0.46)	0.000
White	0.47 (0.50)	0.35 (0.48)	0.000
Married	0.20 (0.40)	0.22 (0.41)	0.933
Female headed family	0.37 (0.48)	0.28 (0.45)	0.025
Homeowner	0.18 (0.38)	0.24 (0.43)	0.000
Previously banked	0.68 (0.47)	0.44 (0.50)	0.000

Figure 3.2 presents the results of the balancing quality checks displaying histogram of the predicted propensity scores for both the treated and control groups. From Figure 3.2, it could be inferred that the propensity score is of equal distribution, suggesting comparability of the treatment and control groups**.** Three matching strategies were used to estimate effects of using prepaid card on financial inclusion as well as AFS use. These matching strategies were one-to-one nearest neighbor matching without replacement, one-to-one nearest neighbor matching with replacement and nearest neighbor caliper matching without replacement.

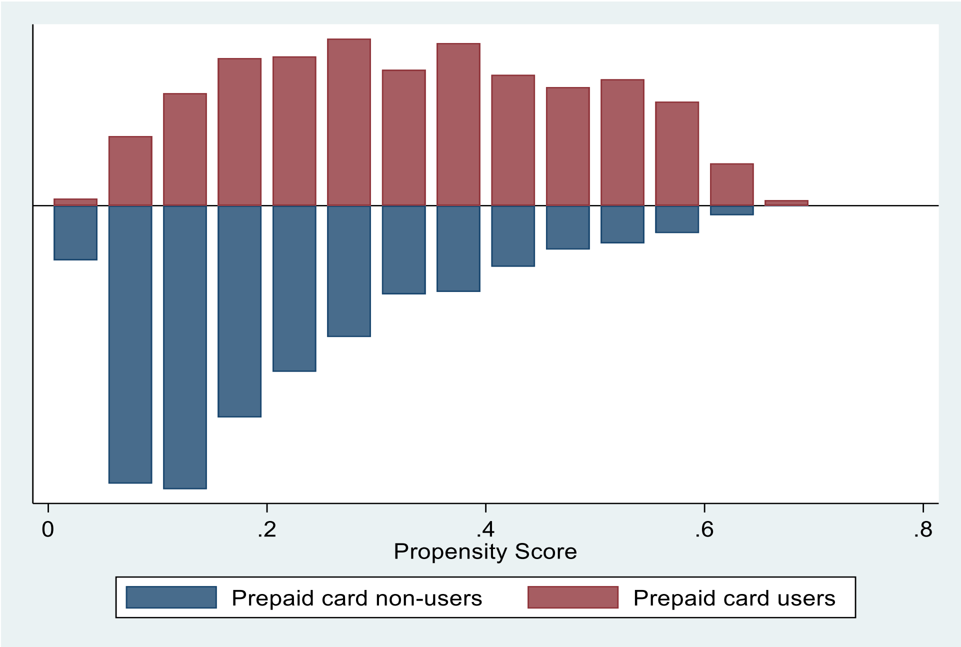

Figure 3.2. Distribution of Propensity scores for the treated and control groups

Tables 3.4 shows the differences in the mean values of the treatment and control groups for the outcome variable; like to open a bank account. The matching process should result in a significant reduction in the bias to make the two groups of users and non-users of prepaid cards comparable ensuring the overall balance of the covariates. The results in tables indicate that the matching process resulted in a significant reduction in the bias and that the treatment and control groups are statistically similar in terms of observable characteristics. The t-values of below 1.96 and corresponding p-values well above the 0.05 threshold for statistical significance at 5% confirm the similarity of the treatment and control groups.

Table 3.4. Mean differences for opening a bank account of matched sample after matching for variables using one to one matching

Variable	Matched (n = 2609)			t - test	
	Treated	Control	Absolute % bias	t	p > t
AFS Credit use	0.4906	0.4729	93.5	1.28	0.202
AFS Transaction use	0.7922	0.7968	97.6	-0.41	0.681
Age	40.522	40.865	92.3	-0.90	0.369
Number of persons in the household	2.7772	2.7126	70.7	1.38	0.169
Spanish only language spoken	0.0323	0.0365	94.4	-0.84	0.403
Rural	0.2382	0.2324	48.1	0.49	0.624
Family income less than $15000	0.5451	0.5551	66.3	-0.72	0.469
High school	0.398	0.3984	98.8	-0.03	0.977
College degree	0.0426	0.0453	-133.9	-0.47	0.636
Employed	0.4433	0.4310	68.8	0.89	0.371
U.S. born / foreign born citizen	0.9343	0.9358	98.8	-0.22	0.822
Black	0.3338	0.3392	84.3	-0.41	0.681
White	0.4687	0.4610	93.3	0.56	0.578
Married	0.1982	0.1832	36.5	1.38	0.169
Female headed family	0.3680	0.3657	97.5	0.17	0.863
Homeowner	0.1821	0.1675	75.3	1.39	0.166
Previously banked	0.6781	0.6750	98.7	0.24	0.813

Figures 3.3 shows graphical density distribution of households who like to open bank accounts after matching for variables. Figure 3.4 shows distribution of propensity scores of the sample before and after matching. The left set of histograms shows propensity score

imbalance in unmatched sample. the right set of histograms shows propensity score balance in the matched sample. The visual inspection of the density distribution of the propensity score in Figure 3.5 indicates a sufficient overlap between the two groups and thus satisfies the required overlap condition of the PSM method. In addition, comparing the density distribution of the propensity scores before and after the matching reveals that the matching process was successful.

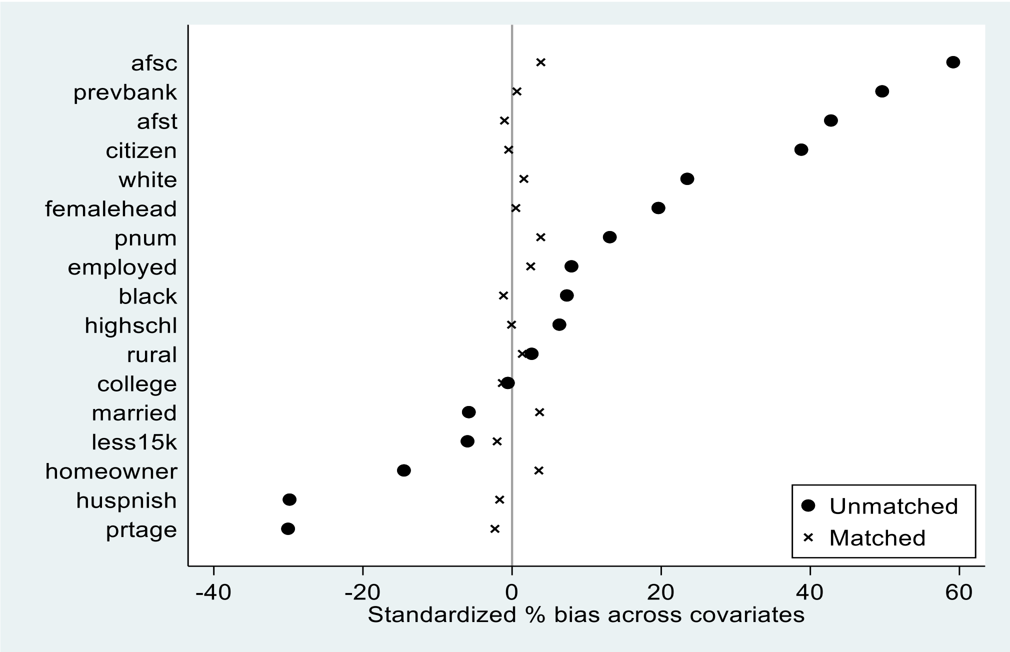

Figure 3.3. Standardized bias plot of households who like to open a bank account of matched sample after matching for variables using nearest neighbor matching

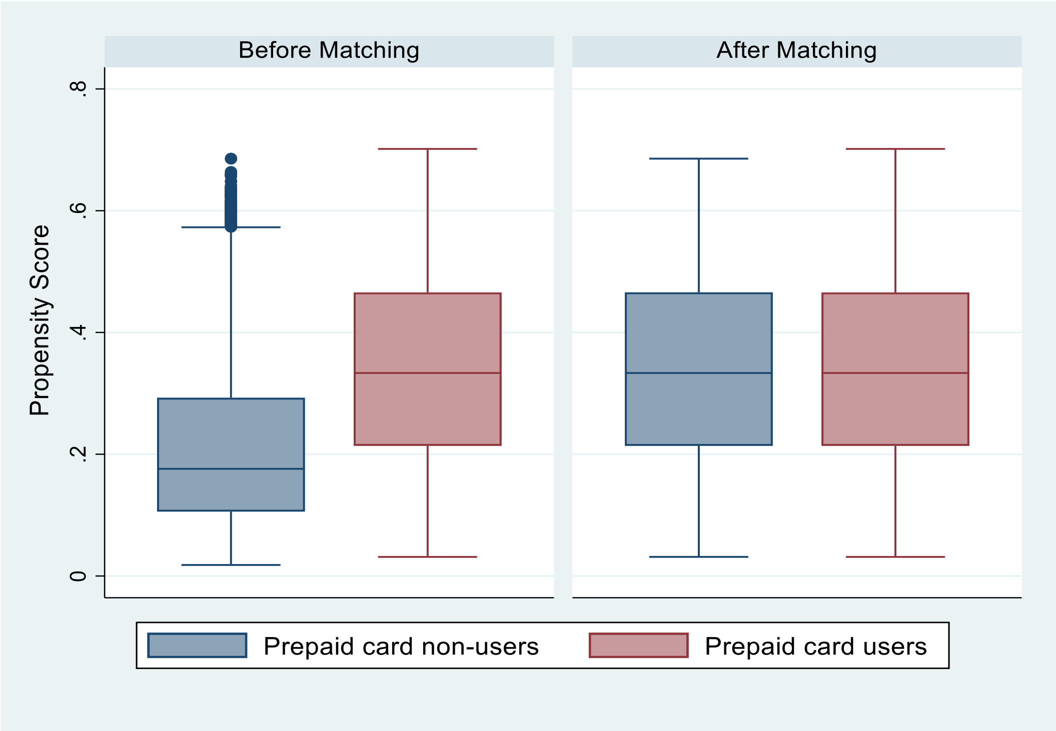

Figure 3.4. Box histogram of propensity Score Balance Before and After Matching

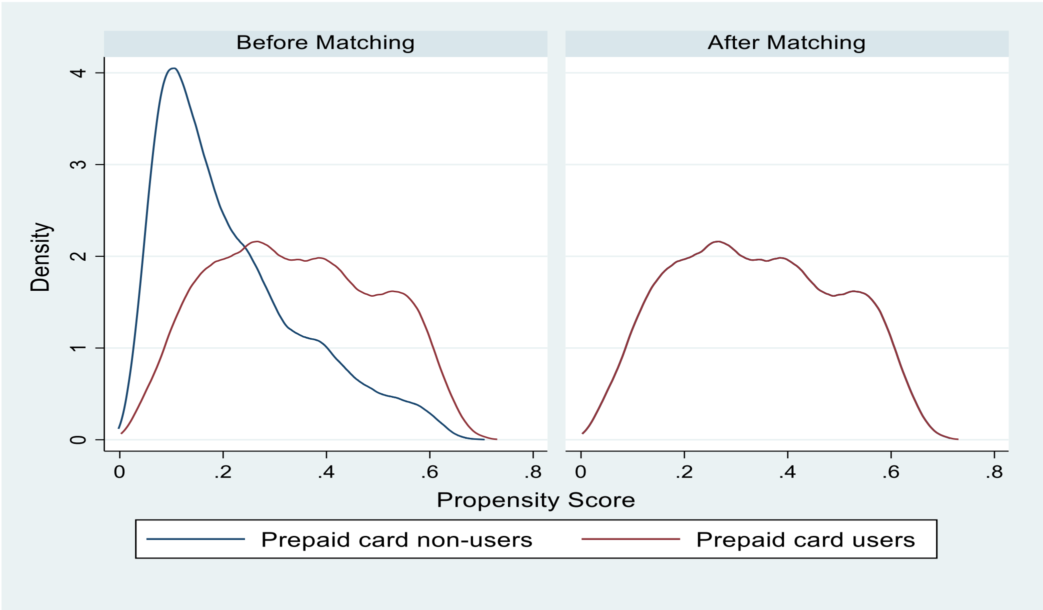

Figure 3.5. Balance density plot of the propensity score before and after matching for treated and control Groups

Also, Figure 3.5 shows that the model used in this study has fulfilled the common support assumption. The intersection of the curve between the prepaid card using households (treatment group) and the prepaid card non-users (control group) before matching and overlap between the two groups after matching represent the same propensity values between the treatment group and the control group.

After matching, standardized bias were computed to assess the matching performance on all covariates that were used to estimate propensity scores. Three matching strategies were used to estimate effects of like to open a bank account on seventeen household indicators. These three matching strategies were one-to-one nearest neighbor matching without replacement, one-to-one nearest neighbor matching with replacement, and one-to-one nearest neighbor matching with caliper. The matching procedures are almost identical for the sample of all prepaid debit card users and the non-prepaid debit card users. The following results discuss the matching results for the full sample.

Table B.3 summarizes standardizes differences and variances before and after matching on 17 covariates for opening a bank account. After matching, standardized biases were expected to move towards 0, while variances were expected to move towards 1. As shown in the table, standardized biases of all covariates except rural and black moved towards 0 after matching, and thirteen covariates showed their variances moving towards 1 after matching except variances of rural and black covariates moved further from 1 after matching.

Table 3.5. Average Treatment effects on treated (ATT) to open a bank account after matching

Matching Method	Coefficient	Standard error	t-stat
Nearest Neighbour 1:1 matching without replacement	-0.0595***	0.0138	-4.33
Nearest Neighbour 1:1 matching with replacement	-0.0507***	0.0167	-3.03
Nearest Neighbour caliper matching with replacement	-0.0504***	0.0167	-3.01

Note: ***, **, *, stand for significance at 1%, 5% and 10% respectively.

Table 3.5 summarizes the average treatment effects of plan to access bank services by using different PSM matching algorithms. The average treatment effect on treated (ATT) is the effect that can be observed in the treated group from entire sample. Columns 2, 3, and 4 report the coefficient values, standard errors, and t statistics. Standard errors are not taken into account to calculate propensity scores. Coefficients are important that show in column 2 and reveal that there are significant negative effects on likelihood to open a bank account as t-statistics are above the value of 1.96. The average treatment effect on the treated groups have -0.0595, -0.0507 and -0.0504 for nearest neighbor one-to-one matching without replacement, nearest neighbor one-to-one matching with replacement, and nearest neighbor matching with caliper meaning between 5 and 6% less likely to open a bank account by the unbanked prepaid card using households than the matched control group of prepaid card not using households. These results are similar to the study was done by Anong and Routh (2021) as the prepaid card users have less intention to join banking services compared to no significant association between prepaid debit card use and banking intention.

As an additional exploration of how the use of prepaid cards impacts AFS credit use and AFS transaction use, estimate ATE with the variables used as the outcome/dependent variables to check any treatment effects controlling for the rest of the predictor variables. Tables B.4 and B.5 show the differences in the mean values of the treatment and control groups for the new outcome variables; AFS credit and transaction use. The results indicate that the matching process resulted in a significant reduction in the bias and that the treatment and control groups are statistically similar in terms of observable characteristics. The t-values of below 1.96 and corresponding p-values well above the 0.05 threshold for statistical significance at 5% confirm the similarity of the treatment and control groups.

Figures 3.6 and 3.7 show the graphical density distribution of households who used AFS credit services and AFS transaction services after matching for variables.

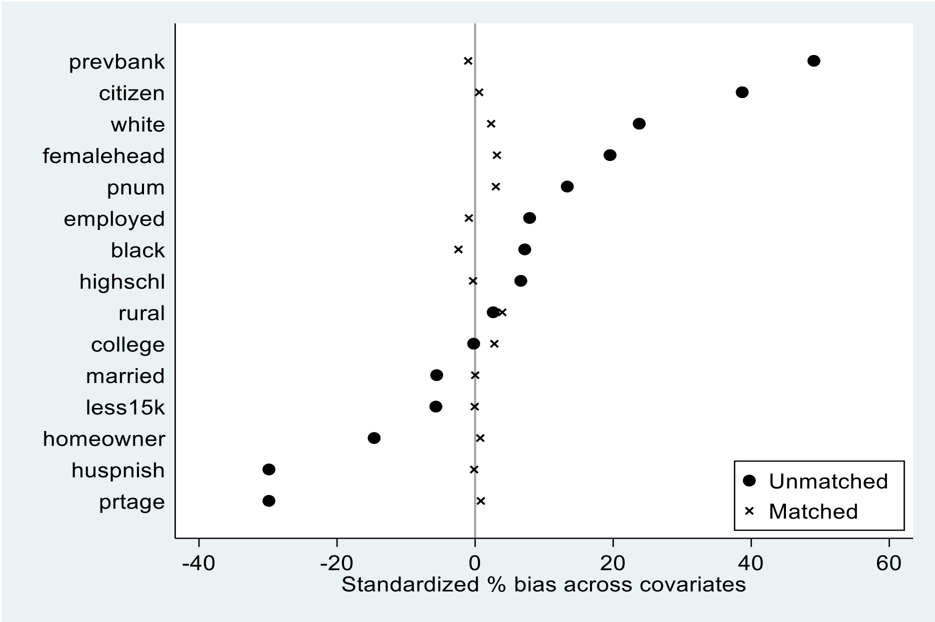

Figure 3.6. Standardized bias plot of households who use AFS credit services of matched sample after matching for variables using nearest neighbor matching

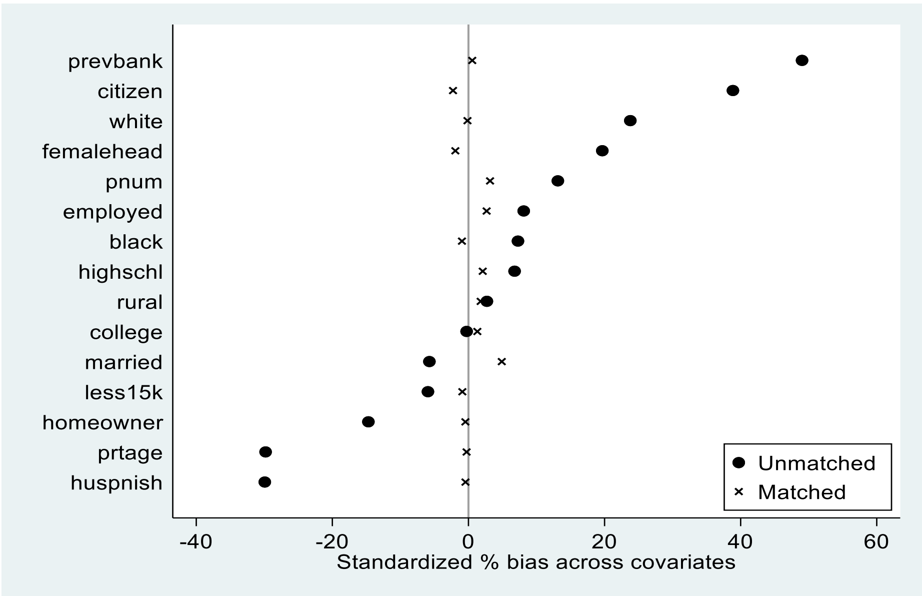

Figure 3.7. Standardized bias plot of households who use AFS transaction services of matched sample after matching for variables using nearest neighbor matching

Table 3.6 presents results yielded from PSM analyses of an average treatment effect for the treated for financial outcomes using nearest neighbor matching without replacement.

Table 3.6. Average Treatment effects on treated (ATT) for the AFS Credit and Transaction outcomes after matching using nearest neighbor matching methods

Outcome variable	One to one matching without replacement			Caliper matching with replacement		
	Coefficient	Standard error	t-stat	Coefficient	Standard error	t-stat
AFS Credit services use	0.1787***	0.0132	13.51	0.1889***	0.0154	12.26
AFS Transaction services use	0.1195***	0.0120	9.94	0.1164***	0.0147	7.93

Notes: ***, **, *, stand for significance at 1%, 5% and 10% respectively.

Results from PSM estimates showed that households using prepaid cards more likely to use AFS credit services by 17.87% and 11.95% use AFS transaction services compared to closely matched households who did not use prepaid cards. The treatment effect of using prepaid cards on all outcomes were statistically significant at the 1% level. Also, columns (4) to (6) summarizes the treatment effects for the outcomes of AFS credit use and AFS transaction use using nearest neighbor matching with caliper with replacement. Standard errors are not taken into account to calculate propensity scores. The output reveals that the average treatment effect on treated (ATT) is the effect that can be observed in the treated group from entire sample. There are significant effects on both outcome variables as t-statistics are above the value of 1.96. The average treatment effect on the treated of 0.1889, meaning 18.89% more use AFS credit services and 11.64% (0.1164) more use AFS transaction services by the unbanked prepaid card using households than the matched control group of prepaid card not using households.

In the tables B.6 in the appendix, standardized biases of all covariates except college degree and black moved towards 0 after matching, and all covariates showed their variances moving towards 1 after matching except variances of college degree and black covariates moved further from 1 after matching. Also, the table B.7, standardized biases of all covariates except rural and black moved towards 0 after matching, and all covariates showed their variances moving towards 1 after matching except variances of rural and black covariates moved further from 1 after matching.

Conclusion

This study examines how the use of prepaid debit cards by unbanked households affects financial inclusion and specifically use potential use formal banking services. Results show that prepaid cards using unbanked households reduced their likelihood to open a bank account by 5%. Also, the study found that unbanked prepaid card users have increased alternative financial credit use by 18%, and transaction services use by 12% than matched unbanked prepaid card non-users. These results reveal that unbanked prepaid debit cards users are less likely to open a bank account. In further investigation, AFS credit and transaction services as outcome variables, results show that unbanked prepaid card users more likely to use alternative financial services as a substitute for traditional banking and positively correlates with using AFS credit services and transaction services. These findings suggests that unbanked households who use prepaid cards have a tendency to use more alternative financial services compared to the more valuable and typically cheaper banking services offered by traditional banks for their financial needs in the household system in US. Overall, the study confirms that unbanked households are open to using financial instruments and services that meet their needs outside of the formal financial system. Nevertheless, unbanked households are yet unable to get all the advantages of formal banking services, and this may explain the persistence of the alternative financial services.

As policy implications, limited access to formal banking services to unbanked households is the main concern moving towards financial inclusion. Thus, prepaid cards provide an opportunity for financial institutions to encourage unbanked community to join formal banking sector and also to benefit financially. With the growth of prepaid card use as a financial tool among unbanked people, offers an unprecedented opportunity for financial institutions to drive loyalty with their current customers and gain the trust from unbanked community and create new revenue streams. Financial institutions should give more attention to the prepaid card industry by promoting prepaid cards offers by their institutions by solving the financial issues of people who are not dealing with formal financial services, and also providing spending control, direct deposit, electronic bill pay and online account access as same as formal bank account services to attract the unbanked community attention towards formal banks.

References

Anong, S. T., & Routh, A. (2021). Prepaid debit cards and banking intention. *International Journal of Bank Marketing*.

Aportela, F. (1999). Effects of Financial Access on Savings by Low-Income People. Mimeo, Banco de México.

Ashraf, Nava, Dean Karlan, and Wesley Yin. (2006). “Deposit Collectors,” *Advances in Economic Analysis & Policy* 6: 1–22.

Barr, Michael S. (2004). “Banking the Poor.” University of Michigan Law School, *Law & Economics Working Papers Archive: 2003–2009*. Available at https:// repository.law.umich.edu /law_econ_archive/art49/

Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine. (2007). “Finance, inequality and the poor.” *Journal of Economic Growth* 12(1): 27-49.

Becker, S. O., and Ichino, A. (2002). Estimation of Average Participation Effects Based on Propensity Scores. The Stata Journal, 2 (4):358-377.

Berry, Christopher. (2004). “To Bank or Not to Bank? A Survey of Low-Income Households.” Joint Center for Housing Studies Working Paper Series, BABC 04-3. February 2004.

Board of governors of the federal reserve system. (2019). Perspectives from Main Street: Bank Branch Access in Rural Communities. Retrieved from: https://www.federalreserve.gov/ publications/ files/bank-branch-access-in-rural-communities.pdf

Brune, Lasse, Xavier Giné, Jessica Goldberg, and Dean Yang. (2011). “Commitments to save: A field experiment in rural Malawi.” *World Bank Policy Research Working Paper Series*, Vol (2011).

Burgess, R. and R. Pande (2005). Can Rural Banks Reduce Poverty? Evidence from the Indian social banking experiment. *American Economic Review* 95(3), 780–795.

Campbell, D., F.A. Martinez-Jerez and P. Tufano (2012). Bouncing Out of the Banking System: An Empirical Analysis of Involuntary Bank Account Closures. *Journal of Banking and Finance* 36, 1224-1235.

Caliendo, M., and S. Kopeinig. (2005). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys* 22: 31–72

Caskey, John P. (2002). “Bringing Unbanked Households into the Banking System.” *Capital Xchange*, January. Available at https://www.brookings.edu/wp-content/uploads/2016/06/ caskey.pdf

Cole, Allison, and Claire Greene. (2016). “Financial Inclusion and Consumer Payment Choice.” Federal Reserve Bank of Boston, *Research Data Reports*, no. 16-05.

Available at https://www.bostonfed.org/publications/research-data-report/2016/financial-inclusion-and-consumer-payment-choice.aspx

Council of Economic Advisors Issue Brief (2016). Financial Inclusion in the United States. Retrieved from: https://obamawhitehouse.archives.gov/sites/default/ files/docs/ 20160610_ financial_inclusion_cea_issue_brief.pdf

Dehejia, R. H., and Wahba, S. (1999). Causal Effects in Non-experimental Studies: Reevaluating the Evaluation of Training Programs. Journal of the American Statistical Association , 94 (448):1053-1062.

Dehejia, R.H., and S. Wahba. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics* 84: 151–61

Demirgüç-Kunt, A. and L. Klapper (2013). Measuring Financial Inclusion: Explaining Variation in Use of Financial Services across and within Countries. Brookings Papers on Economic Activity Spring 2013, 279-340

Dunham, I. M. (2019). Landscapes of financial exclusion: Alternative financial service providers and the dual financial service delivery system. *Business and Society Review*, *124*(3), 365-383.

Dupas, P. and J. Robinson. (2013). "Why Don't the Poor Save More? Evidence from Health Savings Experiments," *American Economic Review*, 103: 1138–71.

Federal Deposit Insurance Corporation. (2021, December 17). How America Banks: Household Use of Banking and Financial Services. Retrieved March 04, 2022, from https://www.fdic.gov/ analysis/household-survey/index.html

Federal Deposit Insurance Corporation (2019). How America Banks: Household Use of Banking and Financial Services, 2019 FDIC Survey (October 2020), from https://www.fdic.gov/ analysis/household-survey/2019report.pdf

Furman, J. (2016). Financial Inclusion in the United States. *Council of Economic Advisers Issue Brief*.

Hogarth, Jeanne M., Christoslav E. Anguelov, and Jinhook Lee. (2005). "Who Has a Bank Account? Exploring Changes Over Time, 1989–2001." *Journal of Family and Economic Issues*, vol. 26, no. 1, pp. 7–30. Available at https://doi. org/10.1007/s10834-004-1410-6

Karlan, Dean and Jonathan Zinman. (2010). "Expanding Credit Access: Using Randomized Supply Decisions to Estimate the Impacts." *Review of Financial Studies* 23(1): 433–64.

Karlan, Dean, Aishwarya Lakshmi Ratan, and Jonathan Zinman. (2014). "Savings by and for the Poor: A Research Review and Agenda." *Review of Income and Wealth* 60(1): 36-78.

Kim, J. H. (2016). A study on the effect of financial inclusion on the relationship between income inequality and economic growth. *Emerging Markets Finance and Trade*, *52*(2), 498-512.

Koenker, R., Bassett, G., (1978). Regression quantiles. Econometrica 33–50.

Le, H. (2017). The Effects of the Dodd-Frank Act on Community Banks.

Morgan, S. L., & Harding, D. J. (2006). Matching estimators of causal effects: Prospects and pitfalls in theory and practice. *Sociological methods & research*, *35*(1), 3-60.

Oh, I., Lee, J. D., Heshmati, A., & Choi, G. G. (2009). Evaluation of credit guarantee policy using propensity score matching. *Small Business Economics*, *33*(3), 335-351.

Osili, U.O. and A. Paulson (2008). What Can We Learn about Financial Access from U.S. Immigrants? The Role of Country of Origin Institutions and Immigrant Beliefs. *World Bank Economic Review* 22(3), 431–455.

Rhine, S.L.W. and W.H. Greene (2006). The Determinants of Being Unbanked for U.S. Immigrants. *Journal of Consumer Affairs* 40(1), 21–40.

Rhine, S.L.W. and W.H. Greene (2013). Factors That Contribute to Becoming Unbanked. *Journal of Consumer Affairs*, 47(1), 27–45.

Rosenbaum, P.R., and D.B. Rubin. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* 70: 41–55.

Rosenbaum, P.R., and D.B. Rubin (1985). "Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score." *American Statistician* 39,1:33–38.

Ruiz Ortega, Claudia. (2013). "From pawn shops to banks: The impact of formal credit on informal households." *World Bank policy research working paper 6634*.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics*, *25*(1), 1.

Toh, Y. L. (2021). Prepaid Cards: An Inadequate Solution for Digital Payments Inclusion. *Federal Reserve Bank of Kansas City, Economic Review*, *106*(4).

Zywicki, T. J. (2013). The Economics and Regulation of Network Branded Prepaid Cards. *Fla. L. Rev.*, *65*, 1477.

Appendix B

Table B.1. Logit regression results of the unbanked prepaid card using and not using households before propensity score matching

Variable	Coefficient	Std. Error	t-test
Constant	-0.219	0.118	-1.62
AFS Credit use	0.086*	0.043	2.68
AFS Transaction use	0.199***	0.001	5.12
Age *(Years)*	-0.003**	0.001	-2.46
Number of persons in the household	0.006	0.015	0.51
Spanish only language spoken *(Yes or no)*	-0.184**	0.083	-2.36
Rural *(Yes or no)*	0.093*	0.049	1.94
Family income less than $15000 *(Yes or no)*	0.098**	0.043	2.08
Highschool diploma *(Yes or no)*	0.016	0.042	0.36
College degree *(Yes or no)*	0.030	0.098	0.40
Employed *(Yes or no)*	-0.009	0.044	-0.15
U.S. born or foreign born citizen *(Yes or no)*	-0.161**	0.071	-2.05
Black *(Yes or no)*	0.029	0.058	0.75
White *(Yes or no)*	-0.064	0.057	-0.79
Married *(Yes or no)*	0.052	0.063	0.80
Female headed family *(Yes or no)*	0.136**	0.054	2.68
Homeowner *(Yes or no)*	0.025	0.049	0.22
Previously banked/ Had a bank account *(Yes or no)*	0.002	0.043	0.73
Number of observations	10633		
LR $chi^2(17)$	80.05		
Prob > chi^2	0.0000		
Pseudo R^2	0.0505		
Log likelihood	- 7275.37		

***, **, *, stand for significance at 1%, 5% and 10% respectively.

Table B.2. Estimation of propensity scores through logistic regression of the unbanked prepaid card using and not using households

Variable	Estimate	Std. Error	t-test
Intercept	-2.479***	0.159	-15.58
AFS Credit use	0.759***	0.051	14.73
AFS Transaction use	0.513***	0.057	8.96
Age *(Years)*	-0.013***	0.002	-7.00
Number of persons in the household	0.069***	0.019	3.71
Spanish only language spoken *(Yes or no)*	-0.330**	0.138	-2.40
Rural *(Yes or no)*	-0.049	0.059	-0.83
Family income less than $15000 *(Yes or no)*	-0.106**	0.053	-2.01
Highschool diploma *(Yes or no)*	-0.021	0.050	-0.43
College degree *(Yes or no)*	0.171	0.121	1.42
Employed *(Yes or no)*	0.069	0.052	1.31
U.S. born or foreign born citizen *(Yes or no)*	0.622***	0.106	5.87
Black *(Yes or no)*	0.324***	0.073	4.45
White *(Yes or no)*	0.419***	0.071	5.94
Married *(Yes or no)*	0.006	0.078	0.08
Female headed family *(Yes or no)*	0.144**	0.064	2.24
Homeowner *(Yes or no)*	-0.273***	0.063	-4.33
Previously banked/ Had a bank account *(Yes or no)*	0.599***	0.052	11.49
Number of observations	10848		
LR $chi^2(17)$	1345.30		
Prob > chi^2	0.0000		
Pseudo R^2	0.1114		
Log likelihood	- 5366.39		

***, **, *, stand for significance at 1%, 5% and 10% respectively.

Table B.3. Summary of Covariate Balance Statistics for opening a bank account

Variable	Standardized Bias Unmatched	Standardized Bias Matched	Variance Unmatched	Variance Matched
AFS Credit use	0.5917	0.0231	1.4611	1.0014
AFS Transaction use	0.4277	-0.0345	0.6860	1.0529
Age	-0.3004	-0.0042	0.7082	0.9492
Household person total	0.1311	0.0243	0.9487	0.8703
Spanish only language spoken	-0.2984	-0.0332	0.3254	0.8454
Rural	0.0263	0.0319	1.0341	1.0414
Family income less than $15000	-0.0598	-0.0270	1.0148	1.0056
Highschool diploma	0.0634	0.0220	1.0314	1.0097
College degree	-0.0056	-0.0094	0.9752	0.9588
Employed	0.0797	0.0023	1.0252	1.0005
U.S. born or foreign born citizen	0.3879	-0.0109	0.3932	1.0397
Black	0.0735	-0.0122	1.0599	0.9916
White	0.2349	0.0100	1.0894	1.0013
Married	-0.0579	0.0223	0.9209	1.0352
Female headed family	0.1963	0.0208	1.1623	1.0119
Homeowner	-0.1449	0.0592	0.8140	1.1092
Previously banked	0.4964	-0.0049	0.8867	1.0038

Table B.4. Mean differences for using AFS credit services of matched sample after matching for variables using nearest neighbour matching

Variable	Matched (n = 2658)			t - test	
	Treated	Control	Absolute % bias	t	p > t
Age	40.45	40.328	97.3	0.32	0.750
Number of persons in the household	2.7833	2.7329	77.6	1.09	0.278
Spanish only language spoken	0.0316	0.0319	99.5	-0.08	0.938
Rural	0.2366	0.2201	-49.9	1.44	0.151
Family income less than $15000	0.5440	0.5444	98.7	-0.03	0.978
High school	0.3999	0.4014	95.3	-0.11	0.911
College degree	0.0429	0.0372	-229.7	1.05	0.294
Employed	0.4458	0.4503	88.4	-0.33	0.741
U.S. born / foreign born citizen	0.9345	0.9327	98.5	0.28	0.783
Black	0.3356	0.3469	66.3	-0.87	0.386
White	0.4688	0.4575	90.3	0.83	0.409
Married	0.1990	0.1990	100.0	0.00	1.00
Female headed family	0.3687	0.3540	83.9	1.11	0.266
Homeowner	0.1821	0.1791	95.0	0.29	0.776
Previously banked	0.6832	0.6881	97.9	-0.38	0.701

Table B.5. Mean differences for using AFS transaction services of matched sample after matching for variables using nearest neighbour matching

Variable	Matched (n = 694)			t - test	
	Treated	Control	Absolute % bias	t	p > t
Age	40.438	44.87	99.1	-0.11	0.913
Number of persons in the household	2.784	2.7307	75.9	1.14	0.255
Spanish only language spoken	0.0316	0.0327	98.5	-0.23	0.816
Rural	0.2370	0.2295	34.2	0.65	0.517
Family income less than $15000	0.5432	0.5477	84.8	-0.33	0.741
High school	0.4008	0.3907	69.2	0.76	0.449
College degree	0.0428	0.0402	-375.3	0.48	0.631
Employed	0.4467	0.4335	67.2	0.97	0.334
U.S. born / foreign born citizen	0.9343	0.9417	94.1	-1.14	0.256
Black	0.3358	0.3403	86.7	-0.35	0.728
White	0.4688	0.4696	99.4	-0.05	0.956
Married	0.1987	0.1788	14.8	1.86	0.063
Female headed family	0.3685	0.3775	90.1	-0.68	0.497
Homeowner	0.1818	0.1837	96.9	-0.18	0.859
Previously banked	0.6829	0.6803	98.9	0.21	0.837

Table B.6. Summary of Covariate Balance Statistics for AFS credit use

Variable	Standardized Bias Unmatched	Standardized Bias Matched	Variance Unmatched	Variance Matched
Age	-0.3016	0.0010	0.6890	0.8899
Household people total	0.2572	0.0251	1.1207	0.9089
Spanish only language spoken	-0.1366	0.0046	0.3653	1.0449
Rural	0.0411	-0.0116	1.0661	0.9831
Family income less than $15000	-0.1242	-0.0585	1.0454	1.0170
Highschool diploma	0.0253	0.0076	1.0092	1.0025
College degree	0.0227	0.0609	1.2106	1.7660
Employed	0.1228	0.0369	1.0527	1.0121
U.S. born or foreign born citizen	0.2023	0.0057	0.3996	0.9670
Black	-0.0135	-0.0468	0.9918	0.9726
White	0.1324	0.0112	1.0120	0.9996
Married	0.0977	0.0493	1.2192	1.0996
Female headed family	0.1614	-0.0441	1.1186	0.9795
Homeowner	-0.0307	0.0236	0.9313	1.0597
Previously banked	0.3677	-0.0231	0.8097	1.0239

Table B.7. Summary of Covariate Balance Statistics for AFS transactions use

Variable	Standardized Bias Unmatched	Standardized Bias Matched	Variance Unmatched	Variance Matched
Age	-0.3017	0.0389	0.6875	0.8758
Household people total	0.2543	0.0242	1.1156	0.8748
Spanish only language spoken	-0.1380	-0.0256	0.3621	0.7953
Rural	0.0427	0.0485	1.0687	1.0785
Family income less than $15000	-0.1267	-0.0499	1.0463	1.0139
Highschool diploma	0.0261	-0.0085	1.0094	0.9973
College degree	0.0231	-0.0149	1.2153	0.8905
Employed	0.1240	-0.0329	1.0529	0.9917
U.S. born or foreign born citizen	0.2040	0.0486	0.3968	0.7666
Black	-0.0121	-0.0756	0.9927	0.9583
White	0.1320	0.0606	1.0118	1.0009
Married	0.0955	0.0532	1.2137	1.1087
Female headed family	0.1639	-0.0058	1.1212	0.9970
Homeowner	-0.0316	0.0683	0.9295	1.1923
Previously banked	0.3675	-0.0146	0.8098	1.0149