

Banking Participation and Income:

Exploring the Possible Correlation Between Bank Account Ownership and Income

Michael K Winder

Southern Utah University

April 2024

Abstract

This paper investigates the potential correlation between owning a checking or savings account and income. The paper uses household data extracted from the June 2021 Current Population Survey (CPS). Utilizing a series of logarithmic regression models, the paper examines whether a relationship exists between banking status and household income while controlling for various demographic and financial factors. The results indicate a significant but diminishing effect of banking participation on income as additional variables are introduced into the models. With these additional control variables, the paper suggests a notable 28.3% increase in income associated with possessing a bank account compared to being unbanked. These results shed light on the intricate dynamics between banking behavior and economic outcomes, offering implications for financial inclusion policies and practices.

Introduction

Financial inclusion, defined as equitable access to banking and financial services, is a cornerstone of economic empowerment and the progress of society. Across diverse international landscapes, initiatives promoting access to basic financial instruments have been very successful. These initiatives have constantly yielded positive results which have catalyzed increased savings, asset accumulation, and enhanced social welfare. A study done in 2010 by Silvia Prina demonstrates the great benefits that can come from having access to financial tools. Her study in Nepal's slums found that offering no-fee basic savings accounts led to increased savings, asset accumulation, and improved household welfare, with 84% of eligible households opening an account and 80% actively using it, resulting in significant increases in monetary and total assets

without crowding out other savings. Such findings underscore the transformative potential of inclusive financial systems in uplifting marginalized communities.

However, despite strides made in financial inclusion, challenges persist, both globally and domestically. Nationally, the United States still has a significant unbanked population numbering in the millions with households lacking financial education and access to financial institutions. While the national unbanked rate has been in decline over the years, there is still a significant unbanked population. Burhouse and colleagues (2016) note that, "The national unbanked rate, measured as the share of U.S. households that do not have a checking or savings account, steadily declined from 8.2 percent in 2011 to 7.0 percent in 2015. Still, 9 million households were unbanked in 2015." This underscores the ongoing need for better access to and education of financial tools.

A common misperception among the American populace is that you need to have a lot of money to have a bank account. This is not true, but the impact of this was seen in a report by the FDIC. According to FDIC (2018), "34 percent of unbanked U.S. households cite 'Do not have enough money to keep in an account' as their main reason for not having a bank account." This statistic is troubling because "having readily accessible sources of cash is of unique importance to life satisfaction, above and beyond raw earnings, investments, or indebtedness."(Ruberton)

This paper endeavors to investigate the correlation between income and banking status in the United States. Utilizing data extracted from the 2021 Current Population Survey, this study uses several logarithmic regression models to explore the relationship between household income and bank account ownership.

Data

To answer my question, “Is there a correlation between income and having a checking or savings account?” I looked at CPS data extracted from the Integrated Public Use Microdata Series (IPUMS) data site. The Current Population Survey (CPS) passes the test as a random sample under the Gauss-Markov assumptions. What makes it pass is due to its rigorous selection process. The CPS uses a multistage, sampling method that starts with the random picking of households and individuals. This ensures that the data collected is a good representation of the entire population, meeting the requirement of a random sample according to the Gauss-Markov assumptions. The CPS team is meticulous about minimizing biases and errors in the survey, adding another layer of credibility to its status as a random sample. When it comes to the other Gauss-Markov assumptions like the independent observations, linear relationships, the absence of perfect multicollinearity, and the notion of homoscedasticity, the CPS data meets these requirements.

I looked at monthly CPS data from June 2021. My research revolves around a binary household variable labeled BUNBANKED. This variable indicates whether someone in a household currently has a checking or savings account. Households that answer "No" are considered "Unbanked" by the FDIC. Because this variable is a household variable I was not able to use individual variables to describe it. This meant that I was unable to use variables such as age, gender, race, or marital status. Instead, I used a series of financial control variables. The following control variables were asked as part of the same survey from June 2021. These variables go into more detail about how the US populace was acquiring its income.

BACCT12M: Anyone in the household has had a checking or savings account in the past 12 months. BMORDER12M: Anyone in the household went to a place other than a bank to purchase a money order in the last 12 months. BPAWN12M: Anyone in the household sold items at a pawn shop because cash was needed in the last 12 months. BRENT2OWN12M:

Anyone in the household had a rent-to-own agreement in the last 12 months.

BTAXLOAN12M: Anyone in the household has taken out a tax refund anticipation loan in the last 12 months. BPAYDAY12M: Used payday loan or payday advance services in the last 12 months. BCREDITCARD: Had a credit card (not debit) from Visa, MasterCard, American Express, or Discover in the last 12 months. HBACCTACCS: The most common way an account was accessed. I also used the variables NMOTHERS and NFATHERS, the number of mothers and fathers in the household, to gain additional household insights.

Analysis of the data revealed several key summary statistics regarding household income and financial behavior. On average, households reported an income of \$78,295.22 annually (see Figure 1.) When examining banked and unbanked income separately, the average banked income stood slightly higher at \$80,206.81 (see Figure 2), while the average unbanked income was notably lower at \$29,650.67 (see Figure 3.) The vast majority of households, approximately 96.22%, possessed a bank account, indicative of widespread banking access. Moreover, 73.43% of households possessed at least one credit card, suggesting a significant portion engaged with credit-based financial services. Conversely, only a small fraction, 1.13%, had utilized payday loans or payday advance services in the past year, indicating limited reliance on these forms of short-term borrowing. Additionally, 9.28% of households had sought alternative avenues, other than banks, to purchase money orders, perhaps reflecting preferences or accessibility constraints. A minimal proportion, 1.01%, resorted to selling items at pawn shops due to immediate cash

needs. Similarly, a small percentage, 1.40%, engaged in rent-to-own agreements, while 1.20% opted for tax refund anticipation loans in the previous year, highlighting alternative financial strategies utilized by a minority of households. It is also worth noting the different ways people accessed their bank accounts in 2021.

The distribution of banking methods among respondents revealed that online banking was the most popular choice, utilized by 43.18% of individuals, followed by mobile banking at 21.30%. ATM or bank kiosk usage accounted for 14.37%, while bank tellers were utilized by 13.18% of respondents. Telephone banking and other methods represented smaller proportions, with 2.44% and 5.54% respectively. You can see this represented graphically in Figure 4. It is interesting to see how more and more people are opting to bank online or on their phones than with an in-person representative. I was able to use this data to run several logarithmic regressions, which will be discussed in the next section.

Method Section

My hypothesis is that there is a significant correlation between having a checking or savings account and income. To assess the possible correlation between banking status and income I use a series of logarithmic models. I use a total of six models, five of which are logarithmic with one being an OLS model. My first model is the base of the five logarithmic models.

Model 1 $\log(hinctot)_h = \beta_0 + \beta_1 bunbanked_h + u$

This model looks at the correlation between banking status and household income. I decided to only have one independent variable so that I could see what the correlation was before adding control variables. Models two through five are as follows.

Model 2

$$\log(hinctot)_h = \beta_0 + \beta_1 bunbanked_h + \beta_2 bcreditcard_h + \beta_3 bpayday12m_h + u$$

Model 3

$$\log(hinctot)_h = \beta_0 + \beta_1 bunbanked_h + \beta_2 bcreditcard_h + \beta_3 bpayday12m_h + \beta'_4 X_h + u$$

Model 4

$$\log(hinctot)_h = \beta_0 + \beta_1 bunbanked_h + \beta_2 bcreditcard_h + \beta_3 bpayday12m_h + \beta'_4 X + \beta'_5 Y + u$$

Model 5

$$\log(hinctot)_h = \beta_0 + \beta_1 bunbanked_h + \beta_2 bcreditcard_h + \beta_3 bpayday12m_h + \beta'_4 X + \beta'_5 Y + \beta'_6 Z + u$$

$$\beta'_4 X = \text{nmothers nfathers}$$

$$\beta'_5 Y = \text{i.bacctaccs}$$

$$\beta'_6 Z = \text{bmother12m bpawn12m brent2own12m btaxloan12m}$$

As you can see, I progressively added financial control variables to the base model. Model 5 is the model that I am using for my analysis of the correlation. I included a series of financial control variables to avoid an omitted variable bias as well. $\beta'_4 X$ contains variables for the number of mothers and fathers in a household. I included these variables because the number of parents in a home can greatly affect household income. $\beta'_5 Y$ contains the variable that reports the most common way a bank account was accessed. This variable was included so that the possible

differences in income could be measured per access method. $\beta'_6 Z$ contains four supplemental financial control variables. These variables represent whether a household has used a money order, pawnshop, rent-to-own agreement, or a tax return anticipation loan in the past 12 months. These four financial services are generally used by households with a lower income, but I am curious to see how much these services affect the total expected household income. My sixth model is as follows.

Model 6

$$hinctot_h = \beta_0 + \beta_1 bunbanked_h + \beta_2 bcreditcard_h + \beta_3 bpayday12m_h + \beta'_4 X + \beta'_5 Y + \beta'_6 Z + u$$

This model is almost exactly the same as model five, but it is an OLS regression instead of a logarithmic one. I decided to use this model in my research to show a nominal approach.

Results

In this section, I will discuss the findings of my models. The results of the six models I used can be seen in Table 1. The results of Model 1 show that the unbanked variable is significant at the 99 percent confidence level. This result shows that those with a bank account can expect to have an income that is 118% higher than the income of someone who doesn't have a bank account. Models two through five have similar results, but the expected effect that the unbanked variable has on household income decreased as more financial control variables were added. The unbanked coefficients for Models 2, 3, and 4 were 0.159, 0.738, and 0.351 respectively.

Model 5 shows that the coefficient of the unbanked variable is 0.283. This means that a household that has a bank account can expect to have a 28.3% higher household income

compared to unbanked households, all else constant. Something that surprised me was that the credit card variable has a positive effect on household income. This could be because people with higher incomes tend to have better credit scores that allow them to have credit cards.

The number of mothers and fathers variables yielded very interesting results. Model 6 shows that a household can expect to earn 39.8% more per year for every additional father in the household. You would think that it would be the same for the number of mothers as well, but it is quite the opposite. For every additional mother, a household can expect to earn 6% less as a household. This result demonstrates a concept called the Motherhood Penalty. “The Motherhood Penalty refers to the phenomenon where women, upon becoming mothers, experience a decline in their career prospects, including lower wages, reduced opportunities for advancement, and increased likelihood of workplace discrimination, stemming from societal and workplace biases towards working mothers.” The number of fathers and mothers variables were both statistically significant at the 99 percent confidence level.

I also want to discuss the financial vehicles that have a negative effect on household income. The payday loan, money order, pawn shop, rent-to-own, and tax return anticipation loan variables were all significant at the 99 percent level. The variables showed an expected decrease in household incomes of 12.7%, 24.1%, 28.3%, 22%, and 25.3% respectively. It is worth noting that these are correlations and not causations. For example, going to a pawn shop and selling something isn't going to decrease your income, but could say that those who need to sell at a pawn shop have a lower income. Model 5 has an R-squared value of 0.274 meaning that the independent variables explain 27.4% of the variation in household income.

I now want to discuss the results of Model 6 which can also be seen in Table 1. This model is a linear version of Model 5. I was curious to see if a linear model would be a better fit for this study. Model 6 has an R-squared value of 0.257, meaning that the independent variables explain 25.7% of the variation in household income. This R-squared value is lower than the one on Model 5, but I still find the results of Model 6 to be important. In this model, results show that a household can expect to earn \$3,405 a year more compared to an unbanked household. In terms of economic significance, \$3,405 a year is not a huge number considering that the average household income is \$78,295.22 annually. The model also shows that a household can expect to earn \$2,436 less per year for every additional mother and earn \$22,508 more for every additional father. This result further demonstrates the effect of the motherhood penalty on a household's income. All of the variables of Model 6 are statistically significant at the 99 percent confidence level with the exception of the unbanked variable at a 95 percent confidence level and telephone banking at the 90 percent confidence level. The results of Models 5 and 6 confirm my hypothesis that there is a correlation between household income and banking status.

Conclusion

This analysis conducted using CPS data from June 2021 has shed light on the correlation between household income and banking status. By employing a series of logarithmic regression models that were supplemented by various financial control variables, the study found that households that have bank accounts tend to have higher incomes compared to unbanked households. The inclusion of additional financial control variables revealed some intriguing insights. The impact of parental status on household income was exemplified by the Motherhood Penalty phenomenon, where households with more mothers tend to earn less income. Additionally, the study found that financial services such as payday loans and pawn shops can

have detrimental effects on a household's income. The results of this study were statistically significant, but more research could be performed to better understand the economic significance of these findings. The findings of this study contribute to a deeper understanding of the complex interactions between household incomes, demographic characteristics, and financial behaviors. The results of this study provide valuable insights for policymakers and financial institutions aiming to address disparities in financial access.

References

Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Megan Schouweiler, and Michael Westberry. IPUMS CPS: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2023.

<https://doi.org/10.18128/D030.V11.0>

Prina, Silvia. "Banking the Poor via Savings Accounts: Evidence from a Field Experiment." *Journal of Development Economics*, North-Holland, 7 Feb. 2015, www.sciencedirect.com/science/article/abs/pii/S0304387815000061.

Hayashi, F., & Minhas, S. (2018). Who are the unbanked? Characteristics beyond income. *Economic Review*, Federal Reserve Bank of Kansas City, 103(2), 55-70.

Shy, O. (2020). Low-income consumers and payment choice. *Research in Economics*, 74(4), 292-300.

Ruberton, P. M., Gladstone, J., & Lyubomirsky, S. (2016, April 11). How Your Bank Balance Buys Happiness: The Importance of "Cash on Hand" to Life Satisfaction. *Emotion*. Advance online publication. <http://dx.doi.org/10.1037/emo0000184>

OpenAI. (2022). Personalized response generated by ChatGPT [AI response].

Figure 1

Household Income

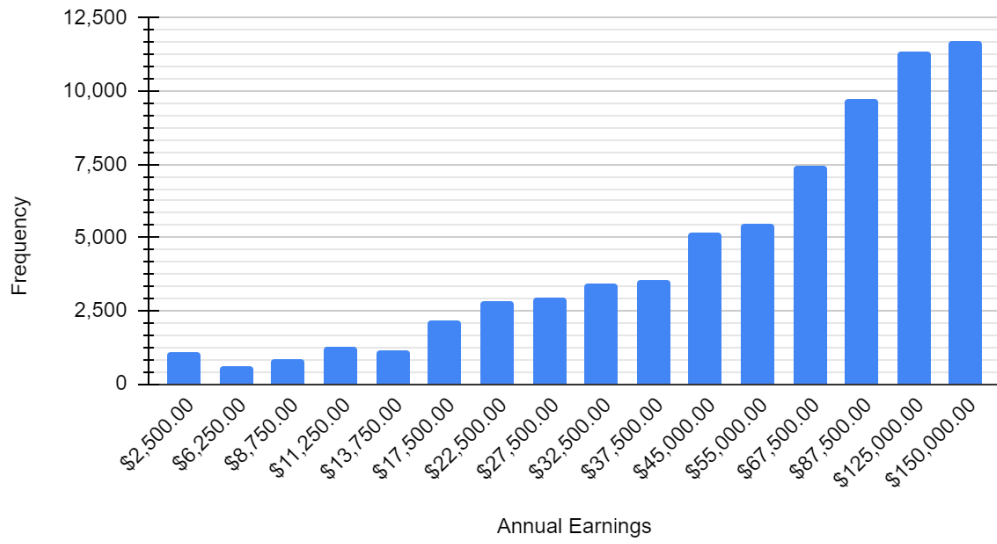


Figure 2

Banked Household Income

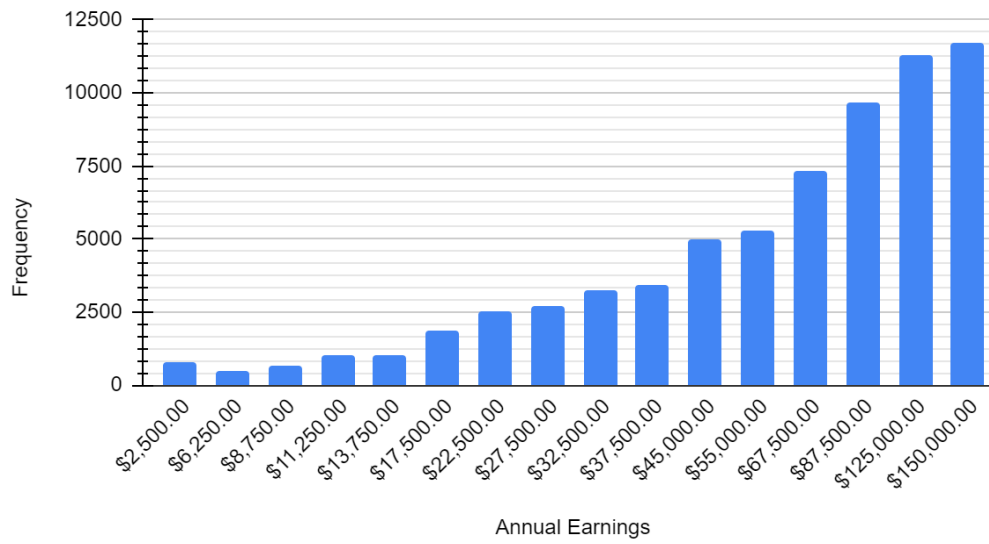


Figure 3

Unbanked Household Income



Figure 4

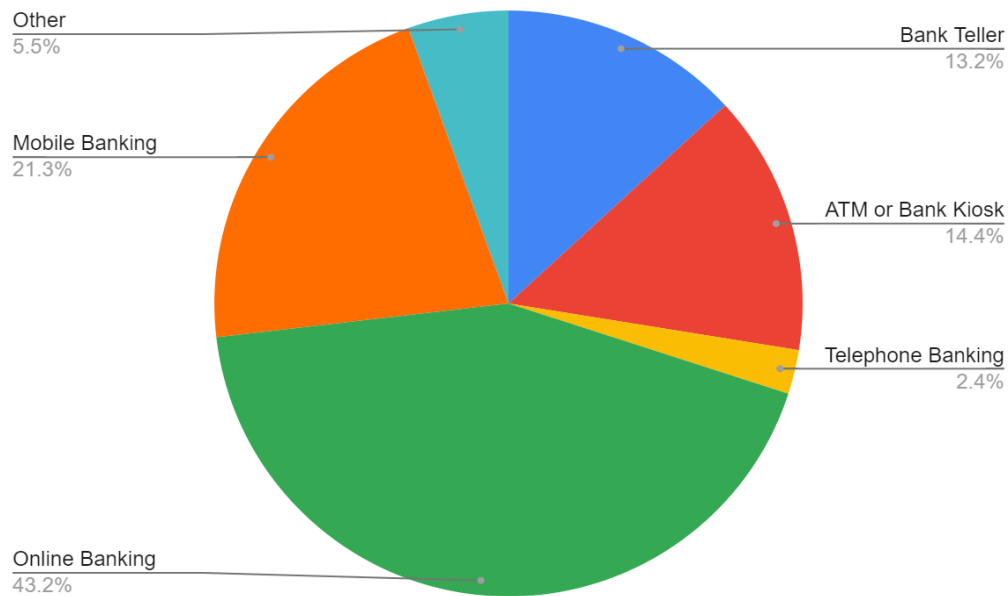


Table 1

Model Number ► VARIABLES	(1) Inhinctot	(2) Inhinctot	(3) Inhinctot	(4) Inhinctot	(5) Inhinctot	(6) hinctot
bunbanked	1.188*** [0.0161]	0.759*** [0.0159]	0.738*** [0.0154]	0.351*** [0.0284]	0.283*** [0.0283]	3,405** [1,590]
bcreditcard		0.627*** [0.00688]	0.598*** [0.00666]	0.518*** [0.00666]	0.492*** [0.00666]	25,918*** [374.2]
bpayday12m		-0.216*** [0.0274]	-0.260*** [0.0265]	-0.252*** [0.0259]	-0.127*** [0.0261]	-11,445*** [1,468]
nmothers			-0.0676*** [0.00673]	-0.0709*** [0.00660]	-0.0610*** [0.00657]	-2,436*** [368.8]
nfathers			0.429*** [0.00737]	0.402*** [0.00722]	0.398*** [0.00717]	22,508*** [402.6]
Bank Teller				0.206*** [0.0257]	0.213*** [0.0255]	7,770*** [1,433]
ATM				0.296*** [0.0256]	0.311*** [0.0254]	13,047*** [1,429]
Telephone Banking				0.0951*** [0.0302]	0.110*** [0.0300]	3,135* [1,683]
Online Banking				0.501*** [0.0250]	0.512*** [0.0249]	26,184*** [1,397]
Mobile Banking				0.648*** [0.0254]	0.645*** [0.0252]	35,314*** [1,417]
Other				0.168*** [0.0455]	0.175*** [0.0452]	9,766*** [2,539]
bmorder12m					-0.241*** [0.00968]	-13,890*** [543.9]
bpawn12m					-0.283*** [0.0277]	-14,341*** [1,554]
brent2own12m					-0.220*** [0.0236]	-14,669*** [1,327]
btaxloan12m					-0.253*** [0.0252]	-15,836*** [1,414]
Constant	9.859*** [0.0158]	9.813*** [0.0150]	9.688*** [0.0148]	9.707*** [0.0144]	9.810*** [0.0147]	26,664*** [825.1]
Observations	70,693	70,693	70,693	70,693	70,693	70,693
R-squared	0.071	0.170	0.228	0.263	0.274	0.257
Standard errors in brackets						
*** p<0.01, ** p<0.05, * p<0.1						