Who Are the Unbanked? Characteristics Beyond Income

By Fumiko Hayashi and Sabrina Minhas

s the U.S. economy recovered from the Great Recession, more households entered the banking system. 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 (Burhouse and others 2016). Still, 9 million households were unbanked in 2015.

Understanding the characteristics of these households is critical in designing effective, tailored policies for financial inclusion. Policymakers and researchers often consider low income to be a defining characteristic of the unbanked. This broad characterization of households, however, may mask large differences in banking status within low-income groups. In particular, low-income households' access to technology, educational attainment, or employment status may also play a role in determining their banking status.

In this article, we conduct a regression analysis to examine which household characteristics beyond income are associated with households' probability of being unbanked. While low-income households have a higher probability of being unbanked on average, we find that the probability of being unbanked varies substantially within this group. Moreover, we find that multiple socioeconomic factors—such

Fumiko Hayashi is a payments policy advisor and economist at the Federal Reserve Bank of Kansas City. Sabrina Minhas is a research associate at the bank. **This article is on the bank's website at www.KansasCityFed.org** as education, age, race, and employment status—as well as technological factors contribute to a low-income household's probability of being unbanked. Of the technological factors we examine, we find that lowincome households without internet access have a much higher probability of being unbanked than those with internet access. Our results suggest that policymakers who promote banking among the unbanked may want to design policies that target low-income households without internet access rather than all low-income households broadly.

Section I describes our data and discusses our empirical methodology for examining households' probability of being unbanked based on household characteristics. Section II confirms that a variety of factors are associated with the probability of being unbanked among households. It then explores the effect of technological access on this probability for low-income households.

I. Data and Empirical Methodology

A household's probability of being unbanked may depend on several different household characteristics. To account for these characteristics, we use data from the 2015 Federal Deposit Insurance Corporation (FDIC) Survey of Unbanked and Underbanked Households.¹ We use the FDIC survey for two reasons. First, the survey enables us to observe each household's banking status alongside detailed household characteristics.² For each household, the survey gathers information on the respondent's sociodemographic characteristics including income, age, race, citizenship, spoken language, education, marital status, employment status, and disability status. In addition, the survey gathers information on each household's technology adoption, such as whether they have internet access or a mobile phone, and each household's geographical characteristics, including their state and proximity to urban areas.

Second, the FDIC survey is highly reliable because its weighted sample is large and nationally representative.³ The FDIC survey is a supplement to the Current Population Survey (CPS), which includes over 60,000 households. Each household in the survey is weighted to adjust for underrepresented populations, which makes the sample nationally representative.

In our analysis, we use the weighted sample, which contains 36,189 households with a non-zero weight.⁴ Table 1 reports summary

Table 1	
---------	--

Summary Statistics for Household Characteristics

Category	Characteristic	Share of sample (percent)	Unbanked rate (percent)
All		100	7.0
Household income	Less than \$15,000	14.1	25.6
	\$15,000 to \$29,999	16.8	11.8
	\$30,000 to \$49,999	19.9	5.0
	\$50,000 or greater	49.2	0.9
Education	Less than high school	10.8	23.2
	High school	55.5	7.5
	College	33.7	1.1
Race	Black Hispanic Other (including white and Asian)	14.1 12.6 73.3	18.2 16.2 3.3
Employment	Employed	61.3	5.0
	Unemployed	3.0	23.0
	Not in labor force	35.7	9.2
Age of household head	34 or younger	21.7	11.2
	35 to 54	35.6	7.8
	55 or older	42.7	4.3
Homeownership	Homeowner	63.3	2.3
	Nonhomeowner	36.7	15.2
Internet access	Has access	72.0	2.7
	No access	28.0	18.2
Mobile phone	Smartphone Feature phone No mobile phone or unknown	67.1 16.7 16.2	4.5 10.7 13.7
Income volatility	Low	88.1	6.3
	High or unknown	11.9	12.8
Citizenship	Citizen	92.8	6.0
	Noncitizen	7.2	20.3
Language	Speaks only Spanish	2.2	31.0
	Speaks other language(s)	97.8	6.5
Marital status	Married	46.7	3.3
	Not married	53.3	10.3
Disability	Disabled Not disabled or not applicable	9.0 91.0	17.6 6.0
Location	Principal city	28.6	10.3
	Suburb or unknown	57.4	5.3
	Rural	14.0	7.6
Region	Northeast	17.8	6.3
	Midwest	21.7	5.7
	South	37.9	8.7
	West	22.6	5.9

Sources: 2015 FDIC survey and authors' calculations.

statistics for the weighted sample, including the share of households with each characteristic and the unbanked rate among a group of households with that characteristic. We draw our statistics from the FDIC statistics reported in Burhouse and others (2016), but we add statistics on internet access and mobile phone ownership and consolidate some characteristics.⁵

Consistent with anecdotal evidence and previous reports, our data show that low-income, less educated, minority-headed, and unemployed households have unbanked rates well above the national average of 7.0 percent. Households with income less than \$15,000 have the highest unbanked rate of 25.6 percent, while households with income \$50,000 or greater have the lowest unbanked rate of 0.9 percent. Those with less than a high school education have the highest unbanked rate of 23.2 percent, while those with a college education have the lowest unbanked rate of 1.1 percent. Unemployed households have an unbanked rate of 23 percent, which is 18 percentage points higher than employed households, who have an unbanked rate of 5 percent. Similarly, black and Hispanic households have unbanked rates of 18.2 percent and 16.2 percent, respectively, while the "Other (including white and Asian)" household group has an unbanked rate of 3.3 percent.

All of the other categories included in Table 1 have at least one characteristic with a higher-than-average unbanked rate. Thus, a wide variety of household characteristics in our sample appear to be related to banking status.

These summary statistics do not show the independent relationship between a particular characteristic and the likelihood of a household with that characteristic being unbanked. Instead, these statistics provide unconditional unbanked rates, which simply denote the share of unbanked households with a given characteristic among all households with that characteristic. This distinction is important. For example, households headed by individuals age 34 or younger and households that do not own a home both have high unbanked rates. Young household heads and homeownership, however, are correlated—that is, households headed by young individuals are less likely to own a home (United States Census Bureau 2018). The summary statistics alone cannot disentangle whether the high unbanked rate among households headed by young individuals is due to age or due to a factor directly related to homeownership. Likewise, while the summary statistics show that low-income households have a higher unbanked rate, they cannot identify whether this high unbanked rate is due to income or another correlated characteristic such as education, employment, or minority status.

To identify which characteristics are *independently* associated with being unbanked for both all households and low-income households, we estimate a binomial probit model using all households in our sample. A binominal probit model is often used to examine how individuals' binary, discrete choices are related to their characteristics by specifying the probability of them choosing one of two alternatives. In our model, individual households choose whether to be unbanked or banked, and household *i*'s probability of being unbanked is regressed on household characteristics. From this model, we are able to estimate the probability of each household being unbanked to see how the distribution of probabilities varies by household characteristic. Furthermore, this model enables us to isolate each characteristic's independent relationship with the probability of being unbanked. The following equation describes the model:

Probability
$$(Y_i = 1) = \Phi(\alpha + \beta X_i),$$

where $Y_i = 1$ if household *i* is unbanked and 0 otherwise, X_i is a vector of household *i*'s characteristics from all 15 characteristic categories observable in our sample (shown in Table 1), and Φ is the cumulative normal distribution function.

Using the coefficients estimated from this model, we calculate marginal effects to quantify the degree to which each characteristic is individually associated with the probability of being unbanked. Specifically, the marginal effect of a given characteristic measures the effect of a change in that characteristic on the probability of being unbanked holding all other characteristics fixed. We use the typical characteristics of our sample as fixed characteristics. Because each of the household characteristics in our model is a binary indicator of whether or not a household has that characteristic, the "typical" characteristics of our sample are the share of the sample with each characteristic.

We calculate two different sets of marginal effects—one for all households and one for low-income households—because certain characteristics may matter more to low-income households than to the full sample. Both sets of marginal effects are calculated using the same coefficients estimated from our model above, but the first set uses the typical characteristics across all households as fixed characteristics while the second set uses the typical characteristics among low-income households as fixed characteristics. The first set of marginal effects allows us to examine the characteristics that change the probability of being unbanked for the entire nationally representative sample. The second set of marginal effects allows us to focus our analysis on low-income households and quantify to what extent characteristics other than income change the probability of being unbanked for households within that group.

II. Household Characteristics Associated with the Probability of Being Unbanked

The results of our probit model estimation show that a variety of household characteristics are independently associated with the probability of being unbanked. Moreover, we find significant variation in the probability of low-income households being unbanked, suggesting income is not the only characteristic associated with households' banking status. Our analysis of only low-income households confirms that a variety of household characteristics are associated with the probability of low-income households being unbanked.

Characteristics associated with the probability of being unbanked for all households

We estimate the probit model and use the estimated coefficients for household characteristics (β) to calculate the marginal effect of each characteristic. Our first set of marginal effects takes all households in our sample into consideration. In other words, we use the typical characteristics across all households when calculating the marginal effects. The marginal effects shown in Table 2 represent the difference between the lowest probability in a category and the probability of a given characteristic in the same category.

Our results are consistent with the summary statistics shown in Table 1. Characteristics in the income, education, race, and employment categories are strongly statistically significant, confirming that these characteristics are associated with an increased probability of being unbanked. However, the degree to which these characteristics are

Table 2 Marginal Effects at the Typical Characteristics across All Households

Characteristic category	Characteristic	Marginal effect	Standard error
Household income	Less than \$15,000 \$15,000 to \$29,999 \$30,000 to \$49,999	0.087*** 0.058*** 0.034***	0.005 0.005 0.005
Education	Less than high school High school	0.062*** 0.037***	0.006 0.005
Race	Black Hispanic	0.046*** 0.034***	0.004 0.004
Employment	Unemployed Not in labor force	0.045*** 0.011***	0.006 0.004
Age of household head	34 or younger 35 to 54	0.054*** 0.048***	0.004 0.004
Homeownership	Nonhomeowner	0.040***	0.003
Internet access	No access	0.048***	0.003
Mobile phone	Feature phone No mobile phone or unknown	0.016*** 0.018***	0.004 0.004
Income volatility	High or unknown	0.010**	0.004
Citizenship	Noncitizen	0.025***	0.005
Language	Speaks only Spanish	0.019***	0.007
Marital status Not married		0.005	0.003
Disability Disabled		0.020***	0.004
Location	Principal city Rural	0.007** 0.008**	0.003 0.004
Region	Northeast Midwest South	0.006 0.010** 0.013***	0.005 0.004 0.004

** Significant at the 5 percent level

*** Significant at the 1 percent level

Notes: We use a weighted probit model. In the probit model, we omit one characteristic from each of the characteristic categories as the controlled characteristic. The omitted characteristic has the lowest unbanked rate among characteristics in the same category.

Sources: 2015 FDIC survey and authors' calculation.

associated with being unbanked varies. The income category has by far the largest marginal effect, suggesting that households with income less than \$15,000 have the highest probability of being unbanked.

The marginal effect of a given characteristic can be interpreted as the difference between the probabilities of being unbanked for two households—one with the given characteristic and one with the characteristic in the same category associated with the lowest probability of being unbanked. For example, the marginal effect of income less than \$15,000 is 0.087, implying the probability of being unbanked for a household with income less than \$15,000 is 8.7 percentage points higher than the probability for a household with income of \$50,000 or greater, after holding all other characteristics fixed at the typical characteristics across all households.⁶

Table 2 shows that the majority of the characteristics in our model are strongly statistically significant even after controlling for income. This suggests that even among low-income households, other factors affect their probability of being unbanked.

The marginal effects shown in Table 2, however, are not well suited to explaining which factors affect low-income households' probability of being unbanked because we use the typical characteristics across all households as the fixed characteristics in calculating those marginal effects. To better understand variations in low-income households' probability of being unbanked, we next examine how the probability distribution varies by income group.

Distribution of household probability of being unbanked by income

We apply the coefficients estimated from the probit model to each household's actual characteristics reported in the survey to obtain each household's individual probability of being unbanked. This allows us to answer questions about differences in banking status among households in the same income group.

As a benchmark, Chart 1 shows how all households in our sample are distributed based on their probability of being unbanked. The horizontal axis shows the probability of being unbanked, which ranges from 0 to 1, while the vertical axis shows the share of households with the probability of being unbanked that is less than or equal to a certain value. The distribution curve for all households is very steep in the low range of probabilities, implying a greater share of households have a low probability of being unbanked. For example, Point A on the curve shows that 80 percent of households have a probability of being unbanked to 0.1. Point B shows that about 90 percent of households have a probability that is less than or equal to 0.25. Thus, about 10 (=90–80) percent of households have a probability of being unbanked between 0.1 and 0.25. More than 90 percent of all households in our sample have a probability less than 0.25, and only 2 percent have a probability greater than 0.5.





The distribution curve for low-income households, however, differs significantly from the curve for all households. Chart 2 plots four different distribution curves for households in each income group. The chart shows that households in the highest income group (those with income of \$50,000 or greater) have a very low probability of being unbanked. Specifically, 99 percent of these households have a probability of being unbanked that is less than 0.1. Households in the lowest and second-lowest income groups (those with income less than \$15,000 and those with income from \$15,000 to \$30,000) have much flatter distribution curves.

Chart 2 has three key implications. First, the vast majority of households in the highest income group have a very low probability of being unbanked regardless of their other characteristics. Second, low income households, especially those with income less than \$15,000, have a higher probability of being unbanked. Third, characteristics other than low income may also be independently associated with the probability of being unbanked. If low income were the only meaningful characteristic, we would expect most low-income households to have a high probability of being unbanked. Thus, we would expect the blue curve for households with an income less than \$15,000 to be flat

over the low probability range and sharply increase over the high probability range. Instead, the smooth curve over almost the entire probability range shows that these households are evenly distributed across the broad range of probabilities. Thus, other characteristics may determine whether low-income households have a higher or lower probability of being unbanked.

Characteristics strongly associated with the probability of being unbanked for low-income households

To identify other characteristics associated with low-income households' probability of being unbanked—and to quantify the extent to which these characteristics change that probability—we calculate another set of marginal effects for households with income less than \$30,000. For this set of marginal effects, we use the same estimated coefficients for household characteristics (β) from our probit model, but use the typical characteristics among households with income less than \$30,000 as the fixed characteristics—in other words, our fixed characteristics are the shares of low-income households with each characteristic in our model. Table 3 shows the marginal effects of characteristics associated with low-income households' probability of being unbanked,

Table 3

Marginal Effects Measured at the Average Characteristics of Low-Income Households: Rank and Magnitude

Rank	Category	Characteristic	Marginal effect	Standard error
1	Education	Less than high school High school	0.143*** 0.085***	0.013 0.011
2	Age of household head	34 or younger 35 to 54	0.123*** 0.109***	0.010 0.009
3	Internet access	No access	0.109***	0.008
4	Race	Black Hispanic	0.106*** 0.078***	0.008 0.010
5	Employment	Unemployed Not in labor force	0.104*** 0.025***	0.014 0.009
6	Homeownership	Nonhomeowner	0.092***	0.007
7	Citizenship	Noncitizen	0.058***	0.011
8	Disability	Disabled	0.047***	0.009
9	Language	Speaks only Spanish	0.044***	0.016
10	Mobile phone ownership	No mobile phone or unkown Feature phone	0.042*** 0.037***	0.010 0.009
11	Region	South Midwest Northeast	0.030*** 0.024** 0.014	0.009 0.010 0.011
12	Income volatility	High or unknown	0.022**	0.010
13	Location	Rural Principal city	0.019** 0.016**	0.009 0.007
14	Marital status	Not married	0.012	0.008

** Significant at the 5 percent level

*** Significant at the 1 percent level

Notes: The omitted characteristic has the lowest unbanked rate among characteristics in the same category. Sources: 2015 FDIC survey and authors' calculations.

ranked from strongest to weakest. As in Table 2, the marginal effects are the difference between the lowest probability in a category and the probability of a given characteristic in the same category.

The results in Table 3 confirm that a variety of factors are associated with low-income households' probability of being unbanked. Interestingly, the magnitudes of all the marginal effects in Table 3 are larger than those in Table 2, suggesting sociodemographic characteristics matter more to the banking status of low-income households than all households. The six categories with the strongest magnitudes—education, age, internet access, race, employment, and homeownership—contain at least one characteristic that is independently and strongly associated with the probability of being unbanked. In each of these categories, the difference between the highest and lowest probabilities is at least 9 percentage points. The next four categories—citizenship, disability, language, and mobile phone ownership—contain characteristics that have a moderate relationship with the probability of being unbanked, with magnitudes in the range of 4 to 7 percentage points. The marginal effects for the remaining four categories—region, income volatility, location, and marital status—have magnitudes smaller than 3 percentage points, suggesting characteristics in these categories are at most weakly associated with the probability.

While we expect categories like education, employment, and race to be associated with low-income households' probability of being unbanked (Burhouse and others 2016), our results for internet access are surprising. Out of the 14 categories for which we control (besides income), internet access has the third-strongest association with low-income households' probability of being unbanked. In addition, Table 3 shows that mobile phone ownership is independently, albeit not strongly, associated with low-income households' banking status. These results may be of interest to policymakers. While interventions based on socioeconomic factors such as age, employment, and race can be difficult and costly to implement, interventions based on technologies such as internet access and mobile phone ownership could be more efficient and effective in promoting banking among low-income households. As such, we discuss our results for technology in greater detail as a potential tool for promoting financial inclusion.

Internet access. Table 3 shows that not having internet access is strongly associated with the probability of being unbanked. A low-income house-hold without internet access has a nearly 11 percentage point greater probability of being unbanked than a low-income household with internet access, even after controlling for all other characteristics.

Despite these controls, internet access may still reflect the influence of other factors. For example, although our model controls for location, three characteristics (rural, suburb, and principal city) may not be sufficient to highlight potential differences in internet access and banking status by geographical location. Rural areas have more limited internet access in general, but both internet access and banking status might vary significantly across rural areas or even within the same rural area.⁷ Examining whether internet access reflects households' geographic location would require further research and more detailed locational information.

Internet access might also reflect households' technology adoption or technological savvy. Households more likely to adopt new technologies might also be more likely to adopt banking services. However, our weaker results for mobile phone ownership, another technology-related category in our model, do not support this interpretation. Instead, the strong relationship between internet access and households' banking status may suggest that online banking is a key channel through which households access banking services.

Mobile phone ownership. The results of our regression analysis show that owning a feature phone, which is a non-smartphone mobile phone, or not owning a mobile phone are independently associated with being unbanked relative to owning a smartphone; however, the strength of this relationship for low-income households is relatively weak. The difference in the probability of being unbanked between a low-income household with a smartphone and a low-income household without a mobile phone is only 4.2 percentage points (Table 3).

The relatively weaker relationship between mobile phone ownership and banking status for low-income households suggests that promoting mobile phone ownership may not be the most effective path to promoting banking services.⁸ This may be a surprising and unwelcome finding for banks and policymakers, many of whom have promoted mobile banking as an access channel to banking services.

III. Conclusion

While the national unbanked rate has steadily declined in recent years, some households still have a relatively high probability of being unbanked. In this article, we use an empirical model to examine which household characteristics are independently and strongly associated with the probability of being unbanked for typical households and for low-income households.

Consistent with anecdotal evidence and existing reports, we find that low income—particularly, income less than \$15,000—has the strongest independent relationship with the probability of being unbanked. However, low-income households exhibit large differences in their probability of being unbanked, and we find a variety of other characteristics play a role in determining these households' banking status.

In addition to socioeconomic characteristics such as education, age, race, and employment, access to technology has an independent association with banking status. In particular, we find that a low-income household without internet access has a significantly higher probability of being unbanked than a low-income household with internet access. Our results suggest that policies that target low-income households without internet access may be able to bring households into the banking system.

Endnotes

¹Unbanked households are households that do not have a checking or savings account, while underbanked households are households that have a checking or savings account but also use alternative financial services outside the banking system, such as check cashing, money remittances, payday loans, and auto title loans.

²Previous studies such as Barr (2002), Caskey (2004), and Hogarth and others (2005) use data from the 1990s or early 2000s and find relationships between the unbanked rate and some sociodemographic characteristics.

³Cole and Greene (2016) use data from the Survey of Consumer Payment Choice (SCPC), conducted by the Federal Reserve Bank of Boston, to examine the relationship between consumers' banking status and their sociodemographic characteristics. While the SCPC is close to nationally representative, its sample size is relatively small. Thus, the data may not have enough variation in terms of unbanked consumers' sociodemographic characteristics.

⁴Households may be allocated a zero weight if they did not respond to the survey or if their demographic characteristics are overrepresented in the survey relative to the national population. We exclude these households to ensure our sample is nationally representative.

⁵We consolidate some characteristics that have similar unbanked rates, such as households with income of \$50,000 or greater, select age groups, and white and Asian households.

⁶The results of probit models are not causal. Characteristics that are significant are associated with households being unbanked, but they do not necessarily cause households to be unbanked.

⁷According to the Federal Communications Commission (2016), 39 percent of rural residents (23 million people) lack high-speed internet access at home compared with 4 percent of urban residents (11 million people). Among rural residents, access to high-speed internet at home varies by state. For example, about 67 percent of rural residents in Alaska lack access, while only 1 percent of rural residents in Connecticut lack access.

⁸Mobile phone ownership and adoption of general purpose reloadable (GPR) prepaid cards are highly correlated, as smartphone owners are more likely to adopt GPR prepaid cards (Hayashi 2016). Thus, promoting mobile phone ownership to the unbanked may also promote access to electronic payment methods.

References

- 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/
- Burhouse, Susan, Karyen Chu, Keith Ernst, Ryan Goodsterin, Alicia Lloro, Gregory Lyons, Joyce Northwood, Yazmin Osaki, Sherrie Rhine, Dhruv Sharma, and Jeffrey Weinstein. 2016. "2015 FDIC National Survey of Unbanked and Underbanked Households." Federal Deposit Insurance Corporation, October. Available at https://www.fdic.gov/householdsurvey/2015/2015report. pdf
- Caskey, John P. 2002. "Bringing Unbanked Households into the Banking System." *Capital Xchange*, January. Available at https://www.brookings.edu/wpcontent/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-datareport/2016/financial-inclusion-and-consumer-payment-choice.aspx
- Federal Communications Commission. 2016. "2016 Broadband Progress Report," January. Available at https://docs.fcc.gov/public/attachments/FCC-16-6A1.pdf
- Hayashi, Fumiko. 2016. "Access to Electronic Payments Systems by Unbanked Consumers." Federal Reserve Bank of Kansas City, *Economic Review*, vol. 101, no. 3, pp. 51–76. Available at https://www.kansascityfed.org/~/media/ files/publicat/econrev/econrevarchive/2016/3q16hayashi.pdf?la=en
- Hayashi, Fumiko, Joshua Hansen, and Jesse Leigh Maniff. 2015. "Driver of Choice? The Cost of Financial Products for Unbanked Consumers," Federal Reserve Bank of Kansas City, Research Working Paper no. 15-15, November. Available at https://www.kansascityfed.org/~/media/files/publicat/reswkpap/ pdf/rwp15-15.pdf
- 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
- Maniff, Jesse Leigh, and W. Blake Marsh. 2017. "Banking on Distributed Ledger Technology: Can It Help Banks Address Financial Inclusion?" Federal Reserve Bank of Kansas City, *Economic Review*, vol. 102, no. 3, pp. 53–77. Available at https://www.kansascityfed.org/~/media/files/publicat/econrev/ec onrevarchive/2017/3q17maniffmarsh.pdf
- United States Census Bureau. 2018. "Quarterly Residential Vacancies and Homeownership, First Quarter 2018." Release No. CB18-57. Available at https:// www.census.gov/housing/hvs/files/currenthvspress.pdf