Online Appendix for

Who reports cryptocurrency to the IRS?

Jeffrey L. Hoopes
Kenan-Flagler Business School
University of North Carolina at Chapel Hill
Chapel Hill, NC

Tyler S. Menzer
Neeley School of Business
Texas Christian University
Fort Worth, TX
tyler.menzer@tcu.edu

Jaron H. Wilde
Tippie College of Business
University of Iowa
Iowa City, IA

Online Appendix
Summary

Online Appendix A: Textual Analysis of IRS Data for Cryptocurrency

We begin our analysis of cryptocurrency transactions by retrieving all transaction records reported to the IRS on Forms 1099-B and 8949 for the years between 2013 and 2021. These records consist of several fields, including a description of the asset, the sales prices, basis, the sales date, the purchase date, and the CUSIP. Not all fields are available for every record. For example, brokers who reported sales on Form 1099-B are only required to report the basis for certain transactions. For Form 8949 records, taxpayers may input these transactions in a variety of ways, for example, sales of similar transactions may be summarized on the form in a single data line.

We begin our search for cryptocurrency transactions with an initial textual search by removing numbers and special characters from transaction descriptions. For Form 1099-B transactions, we remove any records that reported a valid CUSIP number, as cryptocurrency does not have CUSIP identifiers. We identified valid CUSIPs as any CUSIP that was the appropriate length and had a valid check digit. This avoids classifying cryptocurrency ETFs and related products as transactions related to direct interests in cryptocurrency. Next, we search each description for cryptocurrency-related terms (Bitcoin, Ethereum, BTC, ETH, crypto*). A sample of the results from this initial search was then each examined and coded as either cryptocurrency related or noncryptocurrency related to create a second dictionary of exclusion words. As one example, we search for the cryptocurrency related terms ETH, which is the symbol for Ethereum, but we then exclude descriptions that include Raytheon and Ethan Allen, as those are not crypto. Both the dictionary of crypto-related terms and the dictionary of exclusion words was created by the authors by manually examining large samples of descriptions in an iterative process. Our textual analysis likely identifies other cryptocurrencies besides Bitcoin and Ethereum, especially ones with similar names, such as Ethereum Classic or Bitcoin Cash. Through this process, we created a second dictionary of exclusion terms designed to remove common stocks and traditional securities from our sample of cryptocurrency transactions. We then searched the descriptions for our exclusion words and removed any transactions that contained an exclusion word.

After completing the textual search, we manually inspected a random sample of 10,000 cryptocurrency transactions in each of our sample years for each of our five search terms (which leads to a total random sample of 334,800) to assess the possibility of false positives. Although we use a dictionary of five key terms in our process and a full random sample would be 450,000 item descriptions, some terms for some years have fewer than 10,000 observations. In those cases where the total observations are fewer than 10,000, we examine the entire sample. Overall, we find a false positive rate of 3.6%. This false positive rate is highest in 2013 and reaches less than 1% in the middle of our sample period before rising slightly to 3% by 2021. This false positive rate is at a transaction level. Therefore, if a taxpayer reports multiple cryptocurrency transactions, we could still classify that person correctly as a reporting cryptocurrency seller, even if one or more of the transactions we identify are false positives.

Through the random sampling process, we also note a few transactions that highlight the difficulty in assessing data from Form 8949 and self-reported descriptions but also the form's benefits. For example, although cryptocurrency is not subject to wash sale rules, some descriptions specifically note a wash sale adjustment for cryptocurrency. Also, transfers of cryptocurrency assets from one exchange to another are not generally taxable, but we see users reporting the transfers on Form 8949. A number of transactions relate to cryptocurrency mining or staking, consumer purchases such as hotel charges or rent using cryptocurrency, and purchases and sales of privacy coins, such as Monero.

We note several limitations of this approach. Although Form 8949 should reflect all sales of capital assets that a taxpayer makes in a year, it is also self-reported and prone to misspellings, abbreviations, and mistakes as well as the fact that taxpayers will sometimes group multiple transactions under a single description. Using Form 1099-B filings allows us to identify some transactions that taxpayers may have summarized on their tax returns or that they may have been missed due to errors on Form 8949 as well as taxpayers who sell cryptocurrency but do not report those sales on their tax returns. However, 1099-B reporting is not required for all cryptocurrency transactions. Through our textual analysis, we seek to reduce the false positive and false negative rates, but our measurements are not error free.

Online Appendix B: Supplemental Tables

1. Return Filing Type

In Table B1, we examine descriptive statistics of *Reporting Crypto Sellers* based on whether taxpayers e-file or paper file their returns. The type of filing may indicate that taxpayers have different issues or different characteristics. Importantly, we are only able to identify *Reporting Crypto Sellers* who paper file their returns through 1099-B reporting, as we are unable to do a textual search of Form 8949 for paper-filed returns. Overall, we find that the statistics are similar between the two groups, although we note several differences. E-Filing returns have generally higher income but are less likely to include itemized deductions (*Sch A*). Despite being less likely to have itemized deductions, they are more likely to reflect home ownership (*Homeowner).* Unsurprisingly, E-Filers are also more likely to have a *Paid Preparer*, which should be expected, as the IRS requires most paid preparers to e-file returns (IRC Section 6011(e)(3)).

2. Descriptive Statistics by Virtual Currency Checkbox

As we discuss in Section 4.3, starting in 2019, the IRS began asking taxpayers to check a box if they had certain cryptocurrency transactions. In Table B2, we divide the sample into three groups, depending on whether we identify direct cryptocurrency sales (*Reporting Crypto Seller)* and whether the taxpayer checks "yes" to the virtual currency checkbox questions (*Reporting Virtual Checkbox*). We see that the reporting of cryptocurrency varies greatly. In particular, only a minority of observations include identifiable cryptocurrency transactions and checks in the virtual currency checkbox. In column (1), we identify over six million taxpayer-years that check the virtual currency checkbox but that do not appear to report any cryptocurrency transactions. This may indicate transactions that our textual analysis does not capture. These may be taxpayers who are trading smaller, lesser-known cryptocurrencies that we do not search for. This could also mean that taxpayers have cryptocurrency transactions that are reported elsewhere on their tax returns, such as under miscellaneous income for cryptocurrency staking income. Finally, it could relate to uncertainty about who was supposed to check the virtual currency checkbox. The initial virtual currency question asked whether taxpayers received or otherwise acquired any interest in a cryptocurrency. These transactions would not necessarily be taxable if a taxpayer simply purchased a cryptocurrency. Perhaps more concerning is that over five million taxpayer-years are identified as reporting cryptocurrency transactions but do not check the virtual currency checkbox. As we report, our false positive rate is 3.6% overall for cryptocurrency identification. This suggests that many taxpayers with cryptocurrency transactions either marked "no" or simply did not answer the virtual currency question. This raises the question of how many taxpayers understand or pay attention to the question.

Examining the characteristics between the three groups, we find some differences. The Column (3) taxpayers appear to have economically less income than the other two groups and are younger. They are less likely to file *Sch A* and more likely to be a *student* and file the *EIC.* This may suggest that these taxpayers are less sophisticated than the other two groups. However, Column (3) taxpayers are also more likely to have a paid preparer, indicating that the misreporting may relate to more than just taxpayer characteristics. Overall, we find that there is evidence of potential misreporting relating to the virtual currency checkbox, although we do not have any evidence that taxpayers are avoiding the tax due on their cryptocurrency transactions. This provides initial evidence that the IRS may need to do more taxpayer education related to cryptocurrency reporting.

3. Descriptive Statistics for Meme Stock Regression

Table B3 reports descriptive statistics for the sample of taxpayer-years used to estimate equation (3), as described in Section 5.3 of the main paper.

Table B1: Descriptive Statistics for Return Filing Type

	E-filing *Reporting Crypto Taxpayer* (n=9,072,985)		**Paper Filing *Reporting Crypto Taxpayer* (n=256,598)**	
Continuous Variables	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Age	34	11	35	13
Wages	78,824	488,400	60,556	120,992
Interest	883	699,405	339	20,749
Dividends	1,351	226,883	982	86,385
Capital Gain	18,642	1,187,556	10,155	863,561
Taxable Income	93,563	1,237,065	65,676	834,089
Indicator Variables				
Sch A	0.1107	0.3138	0.1598	0.3665
Married	0.3879	0.4873	0.3534	0.4780
Single Male	0.4987	0.5000	0.5362	0.4987
Student	0.1810	0.3850	0.2088	0.4065
Indirect Crypto	0.0158	0.1247	0.0147	0.1204
EIC	0.0777	0.2676	0.0651	0.2466
Homeowner	0.3424	0.4745	0.2852	0.4515
Gambler	0.0192	0.1372	0.0225	0.1483
Paid Preparer	0.4564	0.4981	0.2643	0.4409

Notes. Table B1 reports descriptive statistics for the *Reporting Crypto Sellers* (2013–2021) split out between those taxpayers whose returns were e-filed and those taxpayer years that were not e-filed. *Capital Gain* has a lower bound equal to the $3,000 capital loss limitation. *Single Male* and *Married* are part of a categorical variable where the baseline is taxpayers who do not file a joint return and are female. Due to missing values for gender and age in the Social Security Administration database, a few values for those amounts are missing.

Table B2: Descriptive Statistics by Virtual Currency Checkbox

	***Reporting Virtual Checkbox* = 1 *Reporting Crypto Seller = 0* (n=6,143,704)**		***Reporting Virtual Checkbox* = 1 *Reporting Crypto Seller = 1* (n=3,675,939)**		***Reporting Virtual Checkbox* = 0 *Reporting Crypto Seller = 1* (n=5,160,929)**	
	(1)		**(2)**		**(3)**	
Continuous Variables	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Age	37	14	35	11	33	11
Wages	89,451	5,052,433	100,114	439,303	60,099	488,347
Interest	906	82,906	803	56,428	751	925,513
Dividends	2,326	190,685	2,017	348,673	540	30,472
Capital Gain	20,640	1,369,828	30,841	1,547,189	6,371	696,579
Taxable Income	109,182	1,506,937	131,012	1,638,243	59,173	703,411
Indicator Variables						
Sch A	0.1217	0.3270	0.1274	0.3334	0.0852	0.2793
Married	0.3761	0.4844	0.4224	0.4939	0.3561	0.4788
Single Male	0.4397	0.4964	0.4893	0.4999	0.5072	0.4999
Student	0.1490	0.3561	0.1690	0.3748	0.1945	0.3959
Indirect Crypto	0.0153	0.1228	0.0254	0.1573	0.0104	0.1015
EIC	0.0634	0.2437	0.0406	0.1973	0.1071	0.3093
Homeowner	0.3625	0.4807	0.3935	0.4885	0.3007	0.4586
Gambler	0.0138	0.1167	0.0163	0.1265	0.0220	0.1467
Paid Preparer	0.3490	0.4767	0.3799	0.4854	0.5013	0.5000

Notes. Table B2 reports descriptive statistics for taxpayer years for which *Reporting Crypto Seller* equals 1 or *Reporting Virtual Checkbox* equals 1 for tax years 2019 and after. We limit the sample to tax years 2019 and after because 2019 was the first year taxpayers were asked the virtual currency checkbox questions on tax returns. Column (1) is taxpayers who check "yes" on the virtual currency checkbox but for whom we do not identify as having a direct ownership in cryptocurrency (*Reporting Crypto Seller*). Column (2) is all taxpayers who both check the virtual currency checkbox and who we identify as reporting a direct sale of cryptocurrency. Column (3) is taxpayers for whom we identify as reporting a direct sale of cryptocurrency but who do not check the virtual currency checkbox "yes." *Capital Gain* has a lower bound equal to the $3,000 capital loss limitation. *Single Male* and *Married* are part of a categorical variable where the baseline is taxpayers who do not file a joint return and are female. Due to missing values for gender and age in the Social Security Administration database, a small number of values for those amounts are missing.

Table B3: Descriptive Statistics for Meme Stock Regression

	Reporting Non-Crypto Investor (n=23,962,201)		*Reporting Crypto Taxpayer* (n=3,525,373)	
Meme Stock Variables	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
CMW Meme Equal Weight	0.0321	0.1418	0.1245	0.2276
CMW Meme Volume Weight	0.0288	0.1192	0.0951	0.1609
HJZZ Meme Equal Weight	0.0450	0.1642	0.1635	0.2491
HJZZ Meme Volume Weight	0.0423	0.1411	0.1362	0.1825
Continuous Variables				
Age	55	19	36	11
Wages	111,808	856,779	104,225	503,517
Interest	2,914	146,697	916	268,564
Dividends	12,740	513,437	2,128	354,298
Capital Gain	71,319	1,868,973	33,381	1,604,984
Taxable Income	235,010	2,182,986	139,347	1,662,124
Indicator Variables				
Sch A	0.2213	0.4152	0.1388	0.3457
Married	0.5690	0.4952	0.4589	0.4983
Single Male	0.2116	0.4084	0.4470	0.4972
Student	0.0625	0.2420	0.1585	0.3652
Indirect Crypto	0.0116	0.1070	0.0310	0.1733
EIC	0.0149	0.1212	0.0432	0.2032
Homeowner	0.0175	0.1312	0.0202	0.1408
Gambler	0.7270	0.4455	0.5212	0.4995
Paid Preparer	0.4727	0.4993	0.4319	0.4953

Note: Table B3 presents descriptive statistics for taxpayer-years used to estimate equation (3). *CMW Meme Volume Weight* is defined as the percentage of stock sale proceeds identified as meme stocks by Chacon, Morillon, and Wang (2022). *CMW Meme Equal Weight* is the percentage of unique stocks sold in a year that are identified as meme stocks by Chacon, Morillon, and Wang (2022). *HJZZ Meme Volume Weight* is defined as the percentage of stock sales proceeds identified as meme stocks by Hu, Jones, Zhang, and Zhang (2023). *HJZZ Meme Equal Weight* is the percentage of unique stocks sold in a year that are identified as meme stocks by Hu, Jones, Zhang, and Zhang (2023). All other variables are defined in Appendix A.

Online Appendix C: Census Characteristics of Reporting Crypto Sellers

Although IRS data allow us to offer unique insights into the characteristics of *Reporting Crypto Sellers*, the data have limitations. In particular, they do not contain proxies for the level of education or race. The popular press has highlighted differences in cryptocurrency ownership by race. A recent survey by Charles Schwab found that 25% of Black Americans owned cryptocurrency compared with only 15% of white Americans.[1] Bitcoin and cryptocurrency have been touted as both an opportunity for the Black community to create and retain wealth (Macheel 2023) but have also been seen as failing black investors (Stewart 2023). Against this backdrop, we examine the racial characteristics of *Reporting Crypto Sellers*.

We use regression analysis to examine differences in cryptocurrency reporting by racial groups, as measured by the racial characteristics of the zip code, after controlling for other relevant zip code-level characteristics, such as age composition, education, and income. We present descriptive statistics for this model in Table C1. We present our results in Table C2 for the following regression model:

$$Perc\ Reporting\ Crypto\ Seller = \alpha + \beta_k Age\ Groups + \beta_l Education\ Groups + \beta_m Racial\ Groups + \beta_1 Ln(Median\ Income) + \varepsilon_t \qquad \text{(C1)}$$

The dependent variable in all of our regressions is *Perc Reporting Crypto Seller*, which is the percentage of taxpayers in a zip code who report sales of direct cryptocurrency in a given year in Columns (1) through (3). Age Groups are three variables for the percentage of the population in the given zip code. We split age into three groups: *Perc Under 24, Perc 25-44*, and *Perc 45-64*, which are the percentages of the population in the given zip code within the given age range, with the population over 64 being the control group. Education is split into two groups: *Perc Some College* and *Perc College Degree*, with the control being the percentage of the population without any college. We include six racial/ethnic groups (*Perc African American*, *Perc Native American*, *Perc Asian*, *Perc Native Hawaiian*, *Perc Hispanic*, and *Perc Other*). We finally include the natural log of the zip code's median income to control for differences in income between zip codes. Due to differences in the magnitude of the variables, we standardize all our variables to facilitate the interpretation of the coefficients. This regression is a version of an ecological regression, and factors correlated with the racial makeup of zip codes that are not controlled for may induce error in our estimates (Voss 2005).

Overall, we find some evidence that racial demographics are associated with the rates of cryptocurrency reporting in ways that have been noted in the press (Lowrey 2022; Macheel 2023; Stewart 2023). Specifically, we find that zip codes with higher incomes have higher reporting in more recent years and that the relation between age and reporting has evolved. We also find that more educated areas have greater reporting. Consistent with media reporting, areas with higher *Perc African American* moved from a negative and significant association in 2019 ($p < 0.01$) to a positive and significant association in 2021 ($p < 0.01$). The coefficients on *Perc Hispanic* and *Perc Asian* also suggest an increase in the positive association, whereas the coefficients on *Perc Native American* and *Perc Native Hawaiian* suggest a more negative association over time. Importantly, however, note that regressions of this type are prone to omitted variable bias, as there are likely many factors that are correlated with the racial makeup of zip codes that we cannot control for,

[1]https://content.schwab.com/web/retail/public/about-schwab/Ariel-Schwab_Black_Investor_Survey_2022_findings.pdf

and as our regressions are at the zip code level, we do not show direct evidence that any particular individual sold cryptocurrency.

Table C1: Descriptive Statistics for Census Data

Variable	Mean	Std. Dev.	Median
Population	10,850	15,098	3,567
Perc Crypto Taxpayers	0.0082	0.0141	0.0022
Ln(Median Income)	10.9492	0.3944	10.9364
Age			
Perc Under 24	0.2408	0.1378	0.2803
Perc 25-44	0.1651	0.1519	0.2166
Perc 45-64	0.4248	0.2984	0.2934
Education			
Perc Some College	0.2086	0.0689	0.2072
Perc College Degree	0.3412	0.1628	0.3107
Race			
Perc Hispanic	0.0996	0.1686	0.0361
Perc African American	0.0780	0.1563	0.0119
Perc Native American	0.0177	0.0912	0.0011
Perc Asian	0.0230	0.0568	0.0041
Perc Native Hawaiian	0.0012	0.0108	0.0000
Perc Other Races	0.0573	0.0761	0.0336

Note: In Table C1, we report descriptive statistics by zip code. We aggregate the number of *Crypto Sellers* by zip code and divide it by the total number of tax returns per zip code to create the variable *Perc Crypto Taxpayers*. We then merge this variable, by zip code, with data from the Census Bureau's American Community five-year Estimates table, which provides demographic information by zip code. We begin by limiting our sample period to 2017 through 2021, since, prior to 2017, most zip codes had zero *Crypto Sellers*. We restrict our sample to zip codes with at least 100 individuals.

Table C2: Demographic Statistics (Census Data)

Dependent Variable: *Perc Reporting Crypto Seller*	(1)	(2)	(3)
Year	2017	2019	2021
Ln(Median Income)	0.001	0.008***	0.23***
Age	(0.001)	(0.002)	(0.009)
Perc Under 24	-0.011***	0.02***	-0.503***
	(0.001)	(0.003)	(0.058)
Perc 25-44	0.026***	0.061***	-0.842***
	(0.002)	(0.004)	(0.032)
Perc 45-64	-0.009**	-0.01	-0.183***
	(0.004)	(0.009)	(0.025)
Education			
Perc Some College	-0.004***	0.013***	0.111***
	(0.001)	(0.002)	(0.007)
Perc College Degree	0.027***	0.054***	0.304***
	(0.001)	(0.002)	(0.009)
Racial Makeup			
Perc Hispanic	0.001	0.012***	0.078***
	(0.000)	(0.001)	(0.012)
Perc African American	-0.001***	-0.001	0.129***
	(0.000)	(0.001)	(0.005)
Perc Native American	0.002***	-0.008***	-0.075***
	(0.000)	(0.001)	(0.004)
Perc Asian	0.018***	0.046***	0.309***
	(0.001)	(0.002)	(0.01)
Perc Native Hawaiian	-0.003***	-0.007***	-0.064***
	(0.001)	(0.001)	(0.014)
Perc Other Races	0.000	0.002	0.082***
	(0.001)	(0.002)	(0.01)
Observations	30,364	30,271	30,012
Adj R-Squared	0.3259	0.2922	0.4869

** $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.*

Note: Table C2 reports the results of equation C1. The dependent variable in all of our regressions is *Perc Reporting Crypto Seller*, which is the percentage of taxpayers in a zip code who report sales of direct cryptocurrency in a given year (*Reporting Crypto Seller*). Standard errors are reported in parentheses.

Online Appendix D: Comparison between External Sources and IRS Data Sources

1. Additional Information on Table 2

Table 2 reports comparisons between external estimates of cryptocurrency users and IRS data. It examines three surveys, one academic paper, and data from a public company. We explain each data source here.

Our first three estimates come from surveys. All three surveys were chosen as they capture a representative U.S. sample of adults and provide detailed information on the exact questions asked and methodology. The first survey was conducted by the U.S. Federal Reserve as part of its annual Economic Wellbeing of U.S. Households in 2021 survey. That survey reports that 12 percent of respondents had done at least one of the following in the last 12 months (1) bought or held cryptocurrency as an investment, (2) used cryptocurrency to buy something or make a purchase, or (3) used cryptocurrency to send money to friends or family. This question is more in line with IRS tax reporting, as it requires that the respondent have taken the action within the past year. One caveat is that it also includes individuals who would not have any tax reporting requirements. If respondents simply hold cryptocurrency or buy it and send it as a gift to family or friends, then they would not have any tax reporting requirements. The only condition that would definitely require tax reporting is response (2), that is, using the cryptocurrency to make a purchase. From the survey, only 1.7 percent of respondents used cryptocurrency to make a purchase, and thus this can be seen as a lower bound for the number of respondents who would have had tax reporting requirements.

The second survey is Pew Research Center's American Trends survey. Sixteen percent of Pew respondents answered "yes" to this question: "Have you yourself ever invested in, traded, or used a cryptocurrency?" This question encompasses all past actions by the respondent and is less tied to tax reporting. Respondents would answer yes if they had bought a cryptocurrency in 2021 and never sold it, or if they had bought in an earlier year.

Finally, we report the results of the Cryptocurrency 2022 survey by YouGov.[2] For the YouGov survey, 12 percent of respondents currently owned cryptocurrency, while an additional 9 percent did not currently own cryptocurrency but had in the past. This provides evidence that 42.9 percent (9/21) of cryptocurrency users no longer owned cryptocurrency, which is important when comparing survey results to tax reporting, as tax reporting requires current cryptocurrency activity to be reportable.

For each survey, we also report the estimated number of individuals in the United States who report cryptocurrency activity. Since all three surveys aim for a representative sample of U.S. adults, we can map the responses onto U.S. census data. The total population recorded in the 2020 U.S. Census was 331.4 million (Blakeslee et. al. 2023). We then subtract the portion of the population that is under 18 (73.1 million) to get an estimate of the number of U.S. adults. We then multiply the number of total U.S. adults with the percentage of cryptocurrency owners reported in each survey.

Our fourth estimate of cryptocurrency users comes from Aiello et al. (2024). Aiello (2024) uses proprietary transaction-level data from bank and credit card companies to identify cryptocurrency users. Specifically, they code an individual as a cryptocurrency investor if the

[2] The YouGov survey is the only survey not conducted in 2021, but instead respondents were surveyed in the middle of 2022.

individual makes a deposit to a cryptocurrency exchange between 2013 and 2023. They note that they likely to undercount cryptocurrency investors, as they can only see transactions through traditional financial institutions, which may exclude direct purchases or mining transactions. They also cannot observe whether the individuals actually buy cryptocurrency, as they only see the deposit, not the transactions within the exchange. There are two important caveats with the Aiello (2024) data. First, they count any individual who transacted, at any time, as a cryptocurrency user. This would include individuals who bought but never sold their cryptocurrency. These individuals may not have any tax reporting requirements. Second, Aiello et al.'s (2024) sample period is also a longer window than our IRS sample. To the extent that they capture individuals who first bought cryptocurrency in 2022–2023, we would not observe those individuals in our IRS data.

Our final comparison comes from the number of active users from Coinbase. Although Coinbase is only one exchange, it is one of the largest centralized exchanges in the United States and, as a public company, regularly issues reports of key metrics relating to its retail users. Coinbase had an estimated 89 million verified users as of 2021 (Coinbase 2022). According to a survey, 63.8 percent of respondents who owned cryptocurrency used Coinbase as of March 2021 (Measure Protocol 2023). To examine how many Coinbase users may have reportable cryptocurrency transactions, we examine a key metric reported by Coinbase: monthly transacting users (MTUs). Coinbase defines MTUs as retail users who "actively or passively transact with one or more products on the platform" (Coinbase 2022). MTUs are calculated each quarter, and the average number of MTUs for each quarter is reported in Coinbase's Forms 10-Q and 10-K. For our estimate, we assume that the average MTUs reported for each quarter of 2021 are non-overlapping (perfectly separate) but that they are perfectly overlapping within quarters, such that the average MTUs per quarter represent unique individuals who traded in that quarter. For the four quarters of 2021, Coinbase reported 6.1 million, 8.8 million, 7.4 million, and 11.4 million MTUs. This gives us a total of 33.7 million users (Coinbase 2021a, 2021b, 2021c, 2022). This estimate likely overstates the number of active users, as it assumes that there are no users who engage with Coinbase in more than one quarter of the year. We then adjust this number to account for non-U.S. users. To proxy for the percentage of MTUs that are from outside the United States, we use Coinbase's percentage of revenue from foreign countries. Coinbase reported 19.1 percent of its revenue in 2021 came from outside the United States (Coinbase 2022). Adjusting the average MTUs for the estimated percentage of foreign users gives an estimate of 27.2 million MTUs in the United States. Finally, we divide 27.2 million by 63.8 percent to adjust the number for Coinbase's U.S. market share (Measure Protocol 2023). This calculation gives an estimate of 42.7 million cryptocurrency users in the United States.

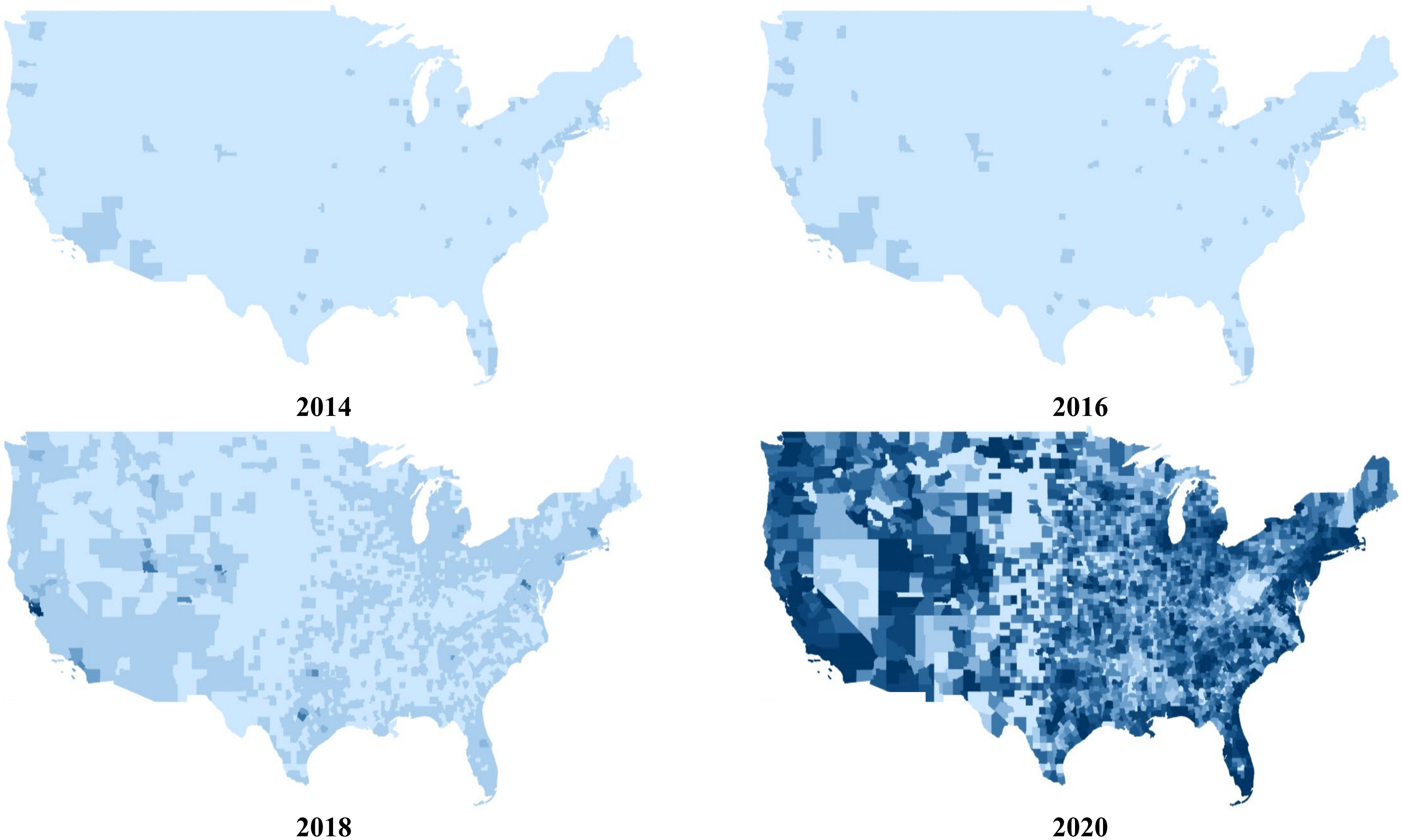
2. Cryptocurrency Reporting by County

In addition to simple percentage comparisons, we also compare geographic distributions over time between IRS data and Aiello et al. (2024). Aiello et al. (2024) report cryptocurrency wealth by county for 2015, 2017, 2019, and 2021. In Figure D1, we examine IRS reporting of cryptocurrency over time for even years in our sample. These are inherently different data: Aiello et al. (2024) are using a dollar-weighted measure of cryptocurrency use, while we examine only the percentage of taxpayers who report any amount of cryptocurrency. Despite this, we believe the comparison can be useful.

In Figure D1, we map the ratio of *Reporting Crypto Seller* tax returns to the total number of tax returns by county for the continental United States for even-numbered years. In the early sample years (2014 and 2016), we see very few counties have cryptocurrency investors, with many counties having none at all.[3] In 2018 and 2020, we observe much broader adoption across the United States, suggesting that cryptocurrency was becoming more geographically widespread. We observe a similar expansion of cryptocurrency reporting to Aiello et al. (2024). In particular, we see higher amounts of cryptocurrency ownership on the West and East Coasts. The reporting of cryptocurrency on tax returns does appear to lag Aiello et al.'s (2024) measure of cryptocurrency wealth, with Aiello et al. (2024) showing nonzero cryptocurrency wealth in many more counties for 2015 than IRS data show for either 2014 or 2016. One striking difference is that, although there is a lack of cryptocurrency reporting in Nevada, West Virginia, and New Hampshire, Aiello et al. (2024) do not show a lack of cryptocurrency wealth. This may indicate that certain states either have very concentrated cryptocurrency wealth or that unreported cryptocurrency income is correlated by state.

[3] Due to restrictions on IRS data and bias in small counties, we set any county with fewer than 10 cryptocurrency reporters or fewer than 1,000 tax returns to 0.

Figure D1: Cryptocurrency Reporting by County



Note: Figure D1 displays a heat map of the percentage of *Crypto Sellers* in each county in the continental United States. Breakpoints between colors are based on the decile rankings for 2020 to make colors comparable between graphs (Breakpoints: 0, >0 to 0.36, 0.36 to 0.45, 0.45 to 0.50, 0.50 to 0.56, 0.56 to 0.62, 0.62 to 0.69, 0.69 to 0.82, 0.82 to 1.01, and 1.01 to 2.30). Some states appear to have low cryptocurrency reporting rates even in 2020. West Virginia, which was rated fifth on a list of the worst states for cryptocurrency investors in 2022 (Newberry 2022) and had the lowest search interest in Bitcoin in 2020 out of all 50 states (Google Trends analysis, untabulated), has a relatively low incidence of cryptocurrency sellers. New Hampshire also has low cryptocurrency reporting and below average Google Trends search volume (rank 41) for 2020. However, Nevada also has low reporting rates but had the highest Google Trend for Bitcoin out of all 50 states in 2020.

Online Appendix E: Analysis of Cryptocurrency Millionaires

One area of interest related to cryptocurrencies is their ability to produce immense wealth as a result of the exponential growth in asset prices. This growth has created rag-to-riches stories for many early investors (Schlott 2022). To examine this phenomenon, we examine individual taxpayers who report large cryptocurrency gains—the cryptocurrency millionaires. To compare the large gains of cryptocurrency millionaires, we also identify two different groups: equity millionaires and gambling millionaires. We define crypto millionaire as a taxpayer who reports gains identified as cryptocurrency of at least \$1 million on either Form 8949 or Form 1099-B in a single year. We define equity millionaire as a taxpayer who reports a capital gain of at least \$1 million and is not a crypto millionaire. Finally, we report a gambling millionaire if the taxpayer's gambling winnings on all Form W-2Gs equals or exceeds \$1 million during the tax year. As we examine both two years before the million-dollar event and two years afterward, we examine only events for the tax years 2015 to 2019. Finally, to examine how unexpected changes in income affect future income, we limit our analysis to only the first time a given taxpayer has \$1 million of a given type of income.

1. Taxable Income Before and After Million Dollar Events

In Figure E1, we report how each group's average Ln(*Taxable Income*) changes both leading up to the \$1 million gain and afterward, separated by *Taxable Income* quartiles in t-2. In each graph, t-0 represents the year of the first million-dollar income event, and correspondingly, we see spikes for all quartiles in the t-0 year, as expected. For all quartiles, the taxpayer's income in t+2 exceeds their income in either t-2 or t-1, indicating that the large increase in income resulted in a persistent increase in income in the future. This effect is most striking for the lowest quartile of income for *Crypto Millionaires.* In t-2, the lowest quartile had an average income of under \$10,000. In the year of the first million-dollar cryptocurrency gain, they have an average income of \$2.1 million and, by t+2 their income, still exceeds \$1 million (untabulated). If we compare average income at t-2 to average income at t+2, the lowest quartile of *Crypto Millionaires* increased average income by 171 times. The middle quartiles 2 and 3 increase income from t-2 to t+2 by 5.7 and 6.4 times respectively, and the top quartile increase their average income by 1.6 times (untabulated).

We also identify 8,836 *Gambling Millionaires* and 250,174 *Equity Millionaires.* We observe similar patterns in both groups compared to *Crypto Millionaires.* We see that the gains from gambling winnings are less persistent than gains from cryptocurrency. The difference between average income at t-2 and t+2 ranges from 61 times in the lowest quartile to 1.2 times in the third and fourth quartiles of income (untabulated). We also observe that the first quartile of *Equity Millionaires* has the lowest persistence out of the three groups, with a ratio of average t-2 income to average t+2 income of only 21.4 times. The persistence for all three upper quartiles of *Equity Millionaire* falls below *Crypto Millionaires.*

2. Portfolio Changes of Cryptocurrency Millionaires

One common perception of cryptocurrency investors is the idea of "HODL," or hold on for dear life, which often signifies investors who will not sell or exit cryptocurrency positions. For *Crypto Millionaires*, there is a question whether they will change their other investment behaviors after such a large recognition of gain or whether they would choose to reinvest their gains in cryptocurrency. To examine the behavior of *Crypto Millionaires* in this regard, we examine *Crypto Millionaire* stock holdings before and after their large cryptocurrency gains.

We report descriptive statistics in Table E1 for the sample of taxpayer-years used to estimate equation (E1). To examine how stock holdings change before and after large gain events, we estimate the following regression:

$$Portfolio\ Characteristic = \alpha + \beta_k Time + \beta_l Type_t + \beta_m Time \times Type_t + \beta_1 Indirect\ Crypto_t + \beta_2 Ln(Interest)_t + \beta_3 Ln(Dividend)_t + \beta_4 Student_t + \beta_5 Homeowner_t + \beta_6 Sch\ A_t + \beta_7 EIC_t + \beta_8 Paid\ Preparer_t + \beta_9 Married_t + \beta_{10} Single\ Male_t + \beta_k Age\ Group_t + \varepsilon_t \tag{E1}$$

Portfolio Characteristic is one of three variables: *Num Unique Stocks*, *Num Trades*, and *Volume*. *Num Unique Stocks* is the number of trades with valid, unique six-digit CUSIPs that a taxpayer sells during the year and proxies for the diversity of taxpayers' portfolios. *Num Trades* is the total number of 1099-Bs received with valid CUSIPs during the year. *Volume* is the total sales price of all stock trades where the taxpayer received a 1099-B with a valid CUSIP.[4] *Time* takes on three values, where *Pre-Million Event* is equal to 1 for the two years before the million-dollar event, *Million Event* is the year the taxpayer has a million-dollar event, and *Post-Million Event* represents the two years afterward. We keep only two observations on either side of the event for each taxpayer. If a taxpayer does not receive any 1099-Bs for a given year, we assume the value of the three dependent variables is 0. *Type* is either *Crypto Millionaire* or *Gambling Millionaire* and defines the type of event for each taxpayer.[5] The coefficient on the interaction term of $Time \times Type$ estimates how the portfolio characteristics of *Crypto Millionaires* and *Gambling Millionaires* evolves compared with the control group of *Equity Millionaires*. All other variables are defined in Appendix A. Our total sample consists of 1,824,872 taxpayer-year observations.

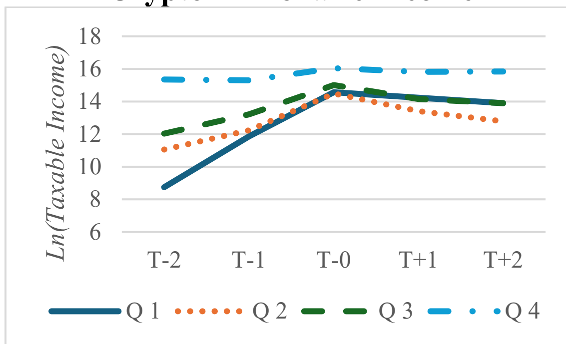
Overall, we find that large, unexpected gains appear to change trading. We present our results in Table E2. We begin by examining *Number Trades* in Column (1). We find that overall there is a large increase in the number of transactions of traditional securities after a large gain, with the difference between *Pre-Million Event* and *Post-Million Event* and *Pre-Million Event×Crypto Millionaire* and *Post-Million Event×Crypto Millionaire* being significantly different at the 1% level. After a large gain, taxpayers make more sales than they did beforehand, even after controlling for taxpayer characteristics. This finding shows that cryptocurrency users take gains from cryptocurrency and diversify into traditional stocks. In Column (2), we examine the number of unique stocks that taxpayers trade. Once again, we find that *Crypto Millionaires* trade more unique stocks after a large gain than before. In Column (3), we examine the volume of sales that taxpayers make in a year. We again find consistent results, with *Crypto Millionaires* selling more securities in the *Post-Million Event* period than in the *Pre-Million Event* period. Our results show that, after large, unexpected gains, *Crypto Millionaires* diversify their investments into regulated securities, providing a pathway for gains from cryptocurrency to affect traditional equity markets.

[4] For all three variables, we transform the raw numbers by adding one and taking the natural logarithm to help with the skewness of the underlying distribution.

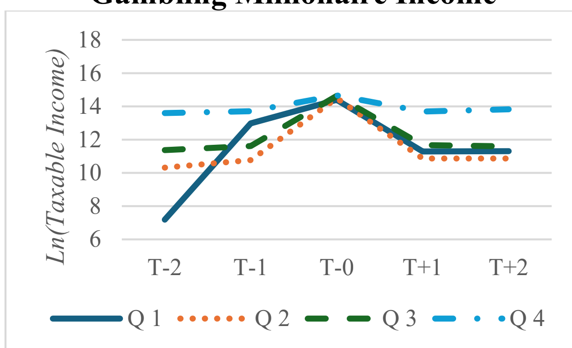
[5] For the regression design, we exclude taxpayers who have more than one type of million-dollar event during the sample period.

Figure E1: Millionaire Income over Time

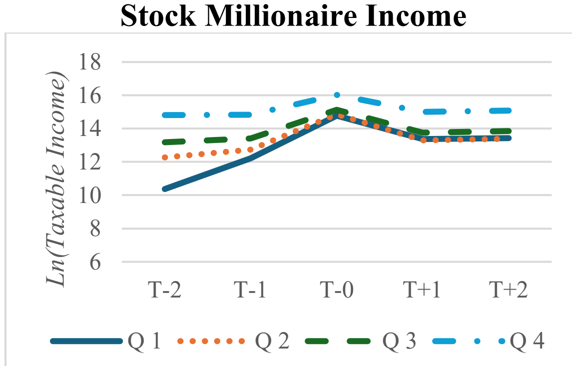
Crypto Millionaire Income



Gambling Millionaire Income



Stock Millionaire Income



Note: Figure E1 shows the average Ln(1+*Taxable Income*) over time of taxpayers around the first "million-dollar event." A million-dollar event is defined as the first time during the sample that a taxpayer experiences income of the related type of at least $1 million. For *Crypto Millionaires*, the income is either gains identified as cryptocurrency from Form 1099-B or gains identified as cryptocurrency from Form 8949. For *Gambling Millionaires*, the income is gambling earnings reported on Form W2-G. For *Equity Millionaires*, the income is capital gain reported on Form 1040, and the person must not have been a *Crypto Millionaire*. Income is shown by quartile (defined in t-2). t-0 is the year of the million-dollar event.

Table E1: Descriptive Statistics for Millionaire Event Regression

Continuous Variables	***Stock Millionaire* (n=1,808,957)**		***Crypto Millionaire* (n=1,167)**		***Gambling Millionaire* (n=14,748)**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Number of Trades	2.53	2.16	3.74	2.19	0.79	1.49
Number Unique Stocks	2.00	1.75	2.85	1.69	0.62	1.16
Volume	9.25	6.14	11.38	5.29	3.24	5.27
Age	58	14	41	11	54	13
Wages	457,695	2,964,793	132,899	643,294	103,386	361,870
Interest	78,902	1,244,249	1,753	8,792	4,485	39,949
Dividends	182,227	3,148,294	5,824	79,100	8,747	84,067
Capital Gain	1,388,020	8,241,650	62,882	173,341	11,787	65,133
Taxable Income	2,262,975	9,538,673	227,690	704,412	1,436,663	8,441,076
Indicator Variables						
Sch A	0.7819	0.4129	0.2494	0.4328	0.7922	0.4057
Married	0.7845	0.4112	0.5621	0.4963	0.5943	0.4911
Single Male	0.1295	0.3357	0.3805	0.4857	0.2469	0.4313
Student	0.0124	0.1105	0.0865	0.2813	0.0157	0.1244
Indirect Crypto	0.0025	0.0496	0.0368	0.1885	0.0005	0.0218
EIC	0.0023	0.0478	0.0471	0.2120	0.0100	0.0997
Homeowner	0.5729	0.4947	0.4704	0.4993	0.5348	0.4988
Gambler	0.0300	0.1707	0.0326	0.1776	1.0000	0.0000
Paid Preparer	0.9196	0.2719	0.5981	0.4905	0.8376	0.3688

Note: Table E1 presents descriptive statistics for taxpayer-years used to estimate equation (E1). *Number Trades* is the natural log of one plus the sum of the number of Form 1099-Bs with valid CUSIPs received by the taxpayer in the year. *Number Unique Stocks* is the natural log of one plus the count of unique six-digit CUSIPs reported for the taxpayer in the year on all forms 1099-B with valid CUSIPs. *Volume* is the Natural Log of one plus the sum of the sales proceeds reported for the taxpayer for the year on all forms 1099-B with valid CUSIPs. *Million Event* is defined as one for the first year in the sample period for which a taxpayer reports over $1 million of income. *Crypto Millionaire* is equal to one if the *Million Event* related to income from cryptocurrency sales reported on either Form 8949 or Form 1099-B. *Gambling Millionaire* is equal to one if the *Million Event* related to income from gambling as reported on Form W2-G. *Equity Millionaires* is equal to one if the taxpayer was not a *Crypto Millionaire* and had *Capital Gain* equal to $1 million or more. All other variables are defined in Appendix A.

Table E2: Regression Analysis of Millionaire Stock Holdings

	(1)		(2)		(3)	
Dependent Variable:	*Number Trades*		*Number Unique Stocks*		*Volume*	
Independent Variables						
Pre-Million Event	-0.0201***		-0.0126***		-0.0922***	
	(0.0029)		(0.0023)		(0.0089)	
Post-Million Event	0.397***	†	0.3244***	†	1.0255***	†
	(0.003)		(0.0024)		(0.009)	
Crypto Millioniare×Pre-Million Event	1.6833***		1.4261***		3.6377***	
	(0.0165)		(0.0127)		(0.0436)	
Crypto Millioniare×Million Event	2.953***		2.37***		7.2652***	
	(0.0619)		(0.0488)		(0.1535)	
Crypto Millioniare×Post-Million Event	2.1848***	†	1.7233***	†	5.1684***	†
	(0.1151)		(0.0902)		(0.2998)	
Gambling Millioniare×Pre-Million Event	0.2547***		0.2548***		0.0291	
	(0.0142)		(0.011)		(0.045)	
Gambling Millioniare×Million Event	0.0053		0.0634***		-0.8011***	
	(0.0114)		(0.009)		(0.0374)	
Gambling Millioniare×Post-Million Event	0.0734***	†	0.0368**	†	0.0474	
	(0.0189)		(0.0147)		(0.0528)	
Observations	1,824,872		1,824,872		1,824,872	
Adj R-Squared	0.299		0.3065		0.3581	

Note: Table E2 presents the regression coefficients from the OLS model from equation (4). *Number Trades* is the natural log of one plus of the sum of the number Form 1099-Bs with valid CUSIPs received by taxpayer in the year. *Number Unique Stocks* is the natural log of one plus the count of unique six-digit CUSIPs reported for the taxpayer in the year on all forms 1099-B with valid CUSIPs. *Volume* is the natural log of one plus the sum of the sales proceeds reported for the taxpayer for the year on all forms 1099-B with valid CUSIPs. *Million Event* is defined as one for the first year in the sample period for which a taxpayer reports over \$1 million of income. *Pre-Million Event* is defined as 1 for the two years preceding the *Million Event. Post-Million Event* is defined as one for the two years after the *Million Event. Crypto Millionaire* is equal to one if the *Million Event* related to income from cryptocurrency sales reported on either Form 8949 or Form 1099-B. *Gambling Millionaire* is equal to one if the *Million Event* related to income from gambling as reported on Form W2-G. The baseline control group for the regression is *Equity Millionaires*, which is equal to one if the taxpayer was not a *Crypto Millionaire* and had *Capital Gain* equal to \$1 million or more. See Appendix A for variables definitions. .* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
†Denotes that the difference between the *Pre- and Post-* variables is significant at the 1% level.

Online Appendix References

Aiello, D., S. Baker, T. Balyuk, M. Di Maggio, M. Johnson, and J. Kotter. 2023. The effects of cryptocurrency wealth on household consumption and investment. NBER Working Paper no. 31445. Cambridge, MA: National Bureau of Economic Research.

Blakeslee, L., Z. Caplan, J. Meyer, M. Rabe, and A. Roberts. 2023. Age and Sex Composition: 2020. 2020 Census Briefs. https://www2.census.gov/library/publications/decennial/2020/census-briefs/c2020br-06.pdf. Accessed September 19, 2025.

Coinbase. 2021a. Form 10-Q Quarterly Report. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001679788/cb843fec-ede4-4ff0-be56-a69062543e57.pdf

Coinbase. 2021b. Form 10-Q Quarterly Report. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001679788/dcec079b-3237-4ae0-a54d-479ded972ab8.pdf

Coinbase. 2021c. Form 10-Q Quarterly Report. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001679788/96565d20-f38b-49e3-a725-0e656b940b09.pdf

Coinbase. 2022. Form 10-K Annual Report https://d18rn0p25nwr6d.cloudfront.net/CIK-0001679788/8e5e0508-da75-434d-9505-cba99fa00147.pdf

Federal Reserve. 2022. Economic Well-Being of U.S. Households in 2021. https://www.federalreserve.gov/publications/files/2021-report-economic-well-being-us-households-202205.pdf. Accessed September 19, 2025.

Lowrey, A. 2022. The Black Investors Who Were Burned By Bitcoin. *The Atlantic.* November 29. https://www.theatlantic.com/ideas/archive/2022/11/black-investors-bitcoin-cryptocurrency-crash/671750/. Accessed December 26, 2023.

Macheel, T. 2023. For the Black community, bitcoin represents an opportunity for wealth preservation. CNBC. February 24. https://www.cnbc.com/2023/02/24/for-the-black-community-bitcoin-represents-an-opportunity-for-wealth-preservation.html. Accessed December. 26, 2023.

Measure Protocol. 2023. Cryptocurrency exchanges used by consumers in the United States from 2021 to 2023. https://www.statista.com/statistics/1288608/crypto-platforms-used-in-usa/. Accessed September 19, 2025.

Pew Research. 2021. American Trends Survey: Politics Timely and Topical. September. https://www.pewresearch.org/wp-content/uploads/2021/11/Crypto-Methods-Topline.pdf. Accessed September 19, 2025.

Schlott, R. 2022. Four Ordinary People Share How They Got Rich from Crypto. *New York Post*. February 5. https://nypost.com/2022/02/05/how-cryptocurrency-made-these-four-ordinary-people-rich/.

Stewart, E. 2023. How Crypto failed Black investors. Vox. February 16. https://www.vox.com/the-goods/23599845/crypto-bitcoin-black-investors-ftx-freedmans-bank-civil-war. Accessed December 26, 2023.

Voss, S. 2005. Aggregation. In *Encyclopedia of Social Measurement*, ed. K. Kempf-Leonard, vol. 1: 33-42. Amsterdam: Elsevier/Academic Press.

YouGov. 2022. Cryptocurrency. July. https://docs.cdn.yougov.com/6gpry6x0mi/toplines_Cryptocurrency.pdf. Accessed September 19, 2025. Accessed September 19, 2025.