United States Government Accountability Office

Report to Congressional Committees

August 2022

TAX EQUITY

Enhanced Evaluation Could Improve Outreach to Small Business Owners

GAO-22-104582

August 2022

TAX EQUITY

Enhanced Evaluation Could Improve Outreach to Small Business Owners

GAO Highlights

Highlights of GAO-22-104582, a report to congressional committees.

Why GAO Did This Study

The COVID-19 pandemic has resulted in significant turmoil in the U.S. economy. Congress enacted tax provisions in pandemic relief efforts to support businesses. However, little is known about the effects of these tax policies by demographic backgrounds of business owners.

The CARES Act includes a provision for GAO to report on its ongoing COVID-19 monitoring and oversight efforts. GAO was also asked to review the effects of selected tax policies on small businesses by race, ethnicity, and sex as part of this oversight.

This report, among other things, estimates use of selected COVID-19 tax provisions by race, ethnicity, and sex of small business owners. It also evaluates potential barriers in accessing COVID-19 tax provisions among small businesses.

GAO analyzed data from IRS, the U.S. Census Bureau, and the Social Security Administration; reviewed literature on analytical methods; and interviewed representatives of small business organizations and agency officials.

What GAO Recommends

GAO is recommending that IRS evaluate its outreach efforts to very small businesses and owners with diverse backgrounds, using relevant and complete information, to inform future outreach. IRS agreed with this recommendation, noting the complexity of evaluating outreach in the absence of demographic data.

View GAO-22-104582. For more information, contact Jessica Lucas-Judy at (202) 512-6806 or lucasjudyj@gao.gov.

What GAO Found

The Internal Revenue Service (IRS) does not collect data on taxpayers, including small business owners, regarding their race, ethnicity, or sex. This makes it difficult to determine if tax provision use varies by demographic group. In the absence of these data, GAO used data from other federal agencies, other taxpayer information, and specific analytical methods to help identify or estimate taxpayers' respective demographic characteristics.

GAO analyzed the use of COVID-19 tax provisions—specifically, paid sick and family leave credits and payroll tax deferrals for employers and the self-employed, as well as the Employee Retention Credit—among a study population of single-owner businesses in tax year 2020. GAO matched data from different agencies to identify the recorded sex of business owners and estimated race and ethnicity of selected taxpayers using an imputation method. This method calculates the probability that a person with a given surname and residential location will identify with selected racial and ethnic groups.

GAO found limited use of the tax provisions by small businesses. Less than 7 percent of eligible small businesses within the study population used the employer and self-employed leave credits or payroll tax deferrals. GAO also found some estimated differential use by demographics of business ownership within the study population. For example,

- **Self-employed leave credits.** GAO estimated that eligible Black or African American- and Hispanic-owned businesses were more likely to use these credits compared to Asian- and White-owned businesses.
- **Employee Retention Credit.** GAO found that a slightly higher percentage of female-owned and Asian-owned businesses used this credit compared to other businesses filing employment tax returns.

Almost all of the small business organizations GAO interviewed identified a poor understanding of the tax provisions as a potential cause of the limited use, particularly among very small businesses. GAO's analysis also identified information and recordkeeping requirements as a potential barrier contributing to limited use. IRS provided information to small businesses on the provisions and used some measures to evaluate its outreach, such as informal feedback and compliance data. However, GAO determined that these measures did not provide relevant and complete information.

A January 2021 Executive Order on advancing racial equity directed agencies to assess their programs and policies to determine if they perpetuate systemic inequalities among groups. Further, the Department of the Treasury's strategic plan includes equity goals involving outreach and education to underserved communities. Enhanced evaluation of ongoing outreach efforts could help IRS develop information useful to groups with different needs, including very small businesses and owners from various demographic backgrounds. While the period of eligibility has passed for these COVID-19 provisions, evaluating outreach could also enhance IRS preparation for communicating tax relief information in future emergencies.

Contents

Figures

Abbreviations

ABS	Annual Business Survey
ACS	American Community Survey
ARPA	American Rescue Plan Act of 2021
BISG	Bayesian Improved Surname Geocoding
CAA	Consolidated Appropriations Act, 2021
EIN	Employer Identification Number
EIP	economic impact payment
EITC	Earned Income Tax Credit
ERC	Employee Retention Credit
FAQ	frequently asked questions
FIPS	Federal Information Processing Standards
IRS	Internal Revenue Service
Leave credits	Paid sick and family leave credits
NES-D	Nonemployer Statistics by Demographics
OMB	Office of Management and Budget
PPP	Paycheck Protection Program
SBA	Small Business Administration
SF1	Summary File 1
SSA	Social Security Administration

U.S. GOVERNMENT ACCOUNTABILITY OFFICE
441 G St. N.W.
Washington, DC 20548

August 3, 2022

Congressional Committees

The COVID-19 pandemic has resulted in significant turmoil with the U.S. economy. Reduced consumer demand early in the pandemic forced both temporary and permanent business closures, particularly among small businesses. Congress enacted several tax provisions in the COVID-19 pandemic relief efforts to help employers support and retain affected employees. This included the Employee Retention Credit (ERC), paid sick and family leave credits (leave credits), and payroll tax deferrals.[1]

As with other tax policies, due to lack of available data, policymakers, researchers, and the public know little about the use of these tax relief provisions by demographics of business owners, such as race, ethnicity, or sex.[2] Disparate use of COVID-19 tax provisions could exacerbate existing inequities in business outcomes among demographic groups. According to a 2020 Federal Reserve Banks survey of small employer businesses, those owned by Asian, Black or African American, and Hispanic individuals experienced more significant negative effects on business revenue, employment, and operations because of the COVID-19 pandemic than White-owned businesses.[3]

We have previously reported on the criteria for a good tax system, which includes equity, among other things.[4] To balance the equity of the U.S. tax system with other policy goals, it is important for policymakers to understand the potential unintended disparities in use of these tax relief

[1]See Families First Coronavirus Response Act, Pub. L. No. 116-127, 134 Stat. 178 (2020); CARES Act, Pub. L. No. 116-136, 134 Stat. 281 (2020).

[2]We refer to "sex" rather than "gender" throughout this report to best reflect the terms used in the datasets we analyzed. Furthermore, our analysis does not consider gender, as it signifies social and cultural factors absent from our data.

[3]Federal Reserve Banks, *Small Business Credit Survey: 2021 Report on Firms Owned by People of Color,* http://www.fedsmallbusiness.org. For purposes of the Small Business Credit Survey, the Federal Reserve Banks define "small employer business" as having one to 499 full- or part-time employees. Small businesses without employees, other than the owner(s), are referred to as "nonemployers."

[4]GAO, *Understanding the Tax Reform Debate: Background, Criteria, & Questions*, GAO-05-1009SP (Washington, D.C.: Sept. 1, 2005). Criteria for a good tax system include considerations of equity, efficiency, simplicity, transparency, and administrability.

provisions by small business owners of varying demographics. Recently, increased social, academic, and media attention on race in the U.S. has renewed policymakers' focus on how our current laws and policies may affect socioeconomic disparities by race and ethnicity. On January 20, 2021, the President signed Executive Order 13985, "Advancing Racial Equity and Support for Underserved Communities Through the Federal Government."[5] The order states that, "A first step to promoting equity in government action is to gather the data necessary to inform that effort."

The CARES Act includes a provision for us to monitor and oversee the federal government's efforts to prepare for, respond to, and recover from the COVID-19 pandemic.[6] In addition, we were asked by the Chairman of the Senate Committee on Finance to review the effects of selected tax policies on small business owners by race, ethnicity, and sex as part of this CARES Act oversight. This report (1) describes the distribution of small business owners by race, ethnicity, and sex; (2) describes information and methods available to examine differences in the use of selected tax provisions by demographics of small business owners; (3) estimates the use of COVID-19 tax provisions by demographics of small business owners for a selected study population; and (4) evaluates potential barriers in accessing COVID-19 tax provisions among small businesses, including whether these barriers varied by demographic group.

To describe the distribution of small business owners by race, ethnicity, and sex, we analyzed data from multiple U.S. Census Bureau datasets. We analyzed 2019 data—the most recent year available—on business characteristics from the Annual Business Survey (ABS) to describe demographics of employer businesses.[7] We also analyzed 2018 data—the most recent year available—from the Nonemployer Statistics by Demographics (NES-D) data series to describe the distribution of demographics for nonemployer small businesses, or businesses with no employees. We used the ABS and NES-D data to examine business demographics by sector. To determine if business ownership was

[5]Exec. Order No. 13985, 86 Fed. Reg. 7009 (Jan. 25, 2021).

[6]Pub. L. No. 116-136, § 19010, 134 Stat. at 579–81.

[7]Due to data suppression issues, we were unable to restrict ABS employer data by demographics to businesses with fewer than 500 employees, a common definition of "small business." Based on our analysis of ABS data, an estimated 99.7 percent of employer businesses have less than 500 employees. Therefore, we use the term "small business" for all ABS employer estimates in this report.

proportional to the demographic characteristics of the population, we analyzed 2019 data from the American Community Survey (ACS). To assess the reliability of the Census datasets, we interviewed knowledgeable agency officials and reviewed technical documentation as well as written responses from Census. We determined that the data used in our analyses were sufficiently reliable for describing the demographic makeup of small business owners.

To identify information for examining differential use of the selected tax provisions by the demographics of business owners, we reviewed federal datasets that include information on the race, ethnicity, and sex of small business owners. We identified datasets for this objective through systematic searches of selected federal agency websites, reviews of relevant research, and interviews with agency officials. We included publicly reported data that selected federal agencies collected in our review. Selected federal agencies included the principal federal statistical agencies, as well as agencies with missions that include collecting data on small businesses.[8] We also reviewed technical documentation that described collection methodology and limitations for each dataset we included in our review.

To identify methods that could be used to obtain missing race, ethnicity, or sex data, we reviewed studies that used or developed methods for this purpose.[9] To identify the studies for our literature review, we searched various databases including Scopus, ProQuest, EBSCO, and Harvard Think Tank Search using search terms related to imputation, estimation, administrative data, race, ethnicity, and sex. We included studies that were published in peer-reviewed journals within the last 10 years; contained original research-based findings; applied a method for obtaining missing race, ethnicity, or sex information for an administrative dataset; and that were methodologically sound. To describe agency

[8]The Office of Management and Budget identified 13 federal statistical agencies, including the Bureaus of Economic Analysis, Justice Statistics, Labor Statistics, and Transportation Statistics; Census Bureau; Department of Agriculture, Economic Research Service; Energy Information Administration; Internal Revenue Service, Statistics of Income; National Agricultural Statistics Service; the National Centers for Education Statistics, and Health Statistics; National Science Foundation, National Center for Science and Engineering Statistics; and Social Security Administration, Office of Research, Evaluation, and Statistics. Selected federal agencies that collect data on small businesses include the Small Business Administration and the Federal Reserve System.

[9]We use the term “missing data” to refer to data that are not measured for some or all records in an administrative dataset, or when survey respondents omit requested information.

efforts to analyze tax provisions by taxpayer demographics, we interviewed officials at Census, the Internal Revenue Service (IRS), and the Department of the Treasury.

To estimate the use of selected COVID-19 tax provisions by race, ethnicity, and sex, we analyzed IRS taxpayer data as of September 2021 to March 2022. Selected tax provisions included the ERC, leave credits for employers and self-employed business owners, and payroll tax deferrals for employers and self-employed individuals. For this analysis, we matched IRS and Social Security Administration (SSA) datasets by Social Security number to obtain the business owners' sex. We also used a prediction method to estimate race and ethnicity. We did this by using IRS and SSA data on business owners' surnames and home addresses as well as publicly available Census data.[10] We then compared how often business owners in each racial and ethnic group used the selected COVID-19 tax provisions for a population of single-owner, eligible businesses. Results of our analysis are not generalizable to the universe of businesses using the tax provisions. We assessed the reliability of Census, IRS, and SSA data by reviewing relevant documentation, interviewing knowledgeable agency officials, and electronic testing. We determined that the data used in our analysis were sufficiently reliable for estimating use of the tax provisions by demographic groups.

To evaluate potential barriers in accessing COVID-19 tax provisions among small businesses, we interviewed 12 organizations that represent or work with small business owners. We selected organizations that were relevant to the groups they represent and advocate for, could speak to broad experiences of small business owners, and were active during the COVID-19 pandemic. We asked representatives of these organizations about owners' experiences using the provisions of interest, including any challenges encountered. We selected organizations that represented a variety of organizational types (e.g., nonprofits and membership organizations) and different demographic groups (e.g., organizations representing Hispanic and women business owners).

[10]Specifically, we used the Bayesian Improved Surname Geocoding (BISG) method. The BISG method uses surname and residential location to predict race and ethnicity for particular demographic groups. Several studies have validated BISG estimates against self-reported racial data collected through administrative or survey sources in the health care and financial sectors. BISG makes estimates by calculating the probability that a person with a given surname and residential location will identify with selected racial and ethnic groups based on data from Census. For additional details about our imputation methodology, see appendix II.

We also reviewed literature on take-up of federal tax provisions to better understand the difficulties taxpayers may face in accessing benefits administered through the tax code. This includes those described in interviews with organizations that represent small business owners. To understand how IRS addressed barriers to accessing tax benefits when administering the COVID-19 provisions, we collected information on its outreach and communication activities. To do this, we reviewed agency documentation and interviewed IRS officials on IRS outreach to small businesses on the COVID-19 provisions. We also collected information on how IRS evaluates its outreach efforts and reviewed relevant IRS policies and procedures. We compared this information to Executive Order 13985, "Advancing Racial Equity and Support for Underserved Communities Through the Federal Government," and the *Treasury 2022-2026 Strategic Plan*.[11] For more information on our scope and methodology, see appendix I.

We conducted this performance audit from October 2020 through August 2022 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Background

U.S. Small Business Population and Federal Tax Requirements

There are roughly 32.5 million small businesses in the United States, representing 99.9 percent of all American businesses, according to the U.S. Small Business Administration (SBA) Office of Advocacy.[12] These small businesses employ almost half of all U.S. workers. A majority—about 26.5 million—of these businesses are nonemployer businesses, meaning they have no employees. Among those businesses that have employees, most employ between 20 and 99 individuals, according to the SBA Office of Advocacy.

[11]Exec. Order No. 13985, 86 Fed. Reg. 7009 (Jan. 25, 2021); Department of the Treasury, *Treasury 2022-2026 Strategic Plan* (Washington, D.C.: 2022).

[12]The specific attributes or thresholds that distinguish small businesses from other firms can vary by industry or government program. The SBA Office of Advocacy defines a small business as an independent business having fewer than 500 employees for research purposes.

Businesses (including small businesses) file specific federal tax forms based on certain attributes of the business, such as the ownership structure and how the business income is taxed. We discuss the following two types of businesses in this report, each with its own required forms and schedules:

- Sole proprietorships (Form 1040, Schedule C *Profit or Loss From Business (Sole Proprietorship*)) are unincorporated and owned by a single individual. Net business income or loss is included in the owner's individual adjusted gross income. Farmers who are sole proprietors use Schedule F (Form 1040, *Profit or Loss From Farming*) to report farm income.
- S corporations (Form 1120-S, U.S. *Income Tax Return for an S Corporation*) cannot have more than 100 shareholders, among other requirements. Income, deduction, loss, and credit recognized at the S corporation level are passed through to its shareholders, who are taxed at the shareholder level.

According to the SBA Office of Advocacy, among businesses without paid employees, approximately 87 percent are sole proprietorships and 5 percent are S corporations. Among small employer businesses, approximately 14 percent are sole proprietorships and 52 percent are S corporations. The remaining third of small employer businesses are either partnerships, C corporations, or other forms of organization, such as nonprofits.[13]

In general, employers must file their employment taxes and report taxable wages and other information. Most employers use Form 941, *Employer's Quarterly Federal Tax Return*, while employers meeting certain criteria may file annually on other forms.[14] In the second through fourth quarters of 2020, approximately 7.2 million employers filed quarterly or annual

[13]Specifically, among small employer businesses, 12 percent are partnerships, 15 percent are corporations, 7 percent are nonprofits, and less than 1 percent are government or other types of businesses. Our focus was sole proprietorships and S corporations because we could identify ownership most easily for these types of businesses through tax returns. We excluded partnerships and C corporations because these types of businesses are more likely to have complex ownership structures and ownership that is difficult to determine based on tax returns. C corporations are owned by shareholders. Corporate income is taxed at the corporate level on taxable income and at the shareholder level on distributed profits.

[14]Annual employment forms include Form 943, *Employer's Annual Federal Tax Return for Agricultural Employees,* and Form 944, *Employer's Annual Federal Tax Return.*

employment tax returns, according to our analysis of IRS data.[15] Employers use employment tax returns to claim employer tax credits or to report payroll tax deferrals. Some business owners are self-employed and generally must pay self-employment tax.[16] Self-employed individuals use income tax returns to claim the selected COVID-19 tax provisions. In 2020, approximately 18.3 million taxpayers filed a self-employment tax form, according to our analysis of IRS data.[17]

COVID-19 Tax Provisions for Small Businesses

During the pandemic, Congress enacted several tax provisions that assist small businesses. Although similarly situated in providing tax relief, the credits and deferrals have different eligibility criteria and other requirements.

- **Employee Retention Credit (ERC).** The ERC encouraged employers to keep employees on the payroll by providing a refundable tax credit based on qualified wages paid to employees, including certain health care expenses. Eligible employers of any size—including tax-exempt entities—could claim the ERC.[18] In 2020, an employer was considered eligible to claim the ERC when it experienced either: (1) full or partial suspension of operations due to governmental orders during any quarter, or (2) significant decline in gross receipts—more than 50 percent for the same quarter in 2019.[19] The Consolidated Appropriations Act, 2021 (CAA), amended aspects of the ERC for credits in 2021, including increased credit maximums, a lower gross receipts threshold, and extending eligibility to previously ineligible employers who received a Paycheck Protection Program (PPP)

[15]Includes employers filing Forms 941, 943, or 944. Also includes employers filing Form 941, Schedule R, and Form 941 amended returns.

[16]According to IRS, generally someone is self-employed if they (1) carry on a trade or business as a sole proprietor or an independent contractor; (2) are a member of a partnership that carries on a trade or business; or (3) are otherwise in business for themselves (including a part-time business).

[17]For self-employed business owners with employees, they can claim the employer provisions on the relevant employment tax return.

[18]Pub. L. No. 116-136, § 2301, 134 Stat. at 347–351; Pub. L. No. 116-260, div. EE, tit. II, §§ 206, 207 134 Stat. 1182, 3059–3065 (2020); Pub. L. No. 117-2, § 9651, 135 Stat. 4, 176–182 (2021).

[19]Employers were no longer eligible in the first quarter after the one in which gross receipts are more than 80 percent of the same quarter in the previous calendar year.

loan.[20] The American Rescue Plan Act of 2021 (ARPA) and the Infrastructure Investment and Jobs Act made additional amendments.[21]

- **Paid sick and family leave credits.** These refundable tax credits were designed to help small employers (with fewer than 500 employees) offset the cost of employee leave related to COVID-19.[22] Qualifying paid sick leave includes quarantine or isolation orders (or caring for someone under orders), seeking a COVID-19 diagnosis, and child care.[23] Qualifying paid family leave includes child care when school or other care is unavailable due to COVID-19. Certain self-employed persons in similar circumstances were allowed equivalent credits.
- **Deferred payroll tax payments for employer share of Social Security tax.** To allow employers to keep additional funds available, the CARES Act granted all employers the option to defer deposits and payments of the employer share of Social Security tax that they would otherwise be required to make during the period beginning March 27

[20]Pub. L. No. 116-260, div. EE, tit. II, §§ 206, 207, 134 Stat. at 3059–3064. PPP loans are made by lenders to small businesses, are guaranteed 100 percent by the Small Business Administration, are low interest, and fully forgivable if certain conditions are met. The eligibility change for PPP loan borrowers was retroactive to 2020. Thus, employers could file an adjusted employment tax return in 2021 to claim the ERC for qualifying wages paid in 2020.

[21]ARPA granted eligibility to "recovery startup businesses" that otherwise would not meet eligibility criteria to claim the credit, among other changes. Pub. L. No. 117-2, § 9651, 135 Stat. at 176–182. The Infrastructure Investment and Jobs Act retroactively terminated the ERC for wages paid after September 30, 2021, for employers other than recovery startup businesses. Pub. L. No. 117-58, § 80604, 135 Stat. 429, 1341 (2021).

[22]Pub. L. No. 116-127, §§ 7001–7004, 134 Stat. 178, 210–219 (2020). The CARES Act provided for advance refunds of the credits and the CAA extended the credits to apply to wages paid before March 31, 2021. Pub. L. No. 116-136, § 3606, 134 Stat. at 411–412; Pub. L. No. 116-260, div. N, tit, II, subtit. B, § 286, 134 Stat. at 1989. Full- and part-time employees are counted. Both credits have maximum payouts.

[23]The tax credits under the Families First Coronavirus Response Act, as amended and extended by the CAA, for leave taken from April 1, 2020, through March 31, 2021, are equal to qualified leave wages paid to employees plus the employer share of Medicare taxes paid with respect to qualified wages and allocable health plan expenses. For periods of leave taken from April 1, 2021 through September 30, 2021, ARPA included codified paid sick and family leave credits that were similar but not identical to the credits enacted under the Families First Coronavirus Response Act. ARPA extended qualifying paid sick leave to vaccinations and time spent recovering from vaccination side effects.

through December 31, 2020.[24] Self-employed individuals could defer half of their Social Security taxes imposed on net earnings from self-employment during the same period. Deferred deposits and payments were to be reported on employment tax returns or income tax returns for self-employed individuals.

In May 2022, we found that employers claimed more than 1.8 million leave credits totaling almost $10 billion for 2020.[25] Additionally, we found that employers claimed 168,918 ERCs totaling about $10.9 billion, for 2020. Employers deferred about $124 billion in payroll taxes, almost all of which were for the employer share of tax deferrals.[26] This prior work also examined the use of these provisions by sector and entity type.

Disparities Exist in Small Business Ownership for Certain Demographic Groups

Ownership Rates

Relative to their share of the U.S. population, Black or African American, Hispanic, and female individuals have estimated lower rates of small business ownership, according to our analysis of Census data (see figs. 1

[24]Pub. L. No. 116-136, § 2302, 134 Stat. at 351–352, as amended by the Paycheck Protection Program Flexibility Act of 2020, Pub. L. No. 116-142, § 4, 134 Stat. 641, 643 (2020). To be considered timely, deferred payments of 50 percent of tax are to be made by December 31, 2021, with the remainder due December 31, 2022. The employer share of Social Security tax is 6.2 percent of taxable earnings up to the Social Security wage base cap on taxable income.

[25]GAO, *COVID-19: IRS Implemented Tax Relief for Employers Quickly, but Could Strengthen Its Compliance Efforts*, GAO-22-104280 (Washington, D.C.: May 17, 2022).

[26]On August 8, 2020, a Presidential Memorandum directed the Secretary of the Treasury to use his authority to allow the deferral of withholding, deposit, and payment of the employee share of certain employment taxes imposed on wages or compensation paid from September 1, 2020, through December 31, 2020. *Deferring Payroll Tax Obligations in Light of the Ongoing COVID-19 Disaster*, 85 Fed. Reg. 49587 (Aug. 13, 2020).

and 2).[27] These disparities in small business ownership are generally greater among employer businesses than nonemployer businesses, meaning they have no employees.[28] For example,

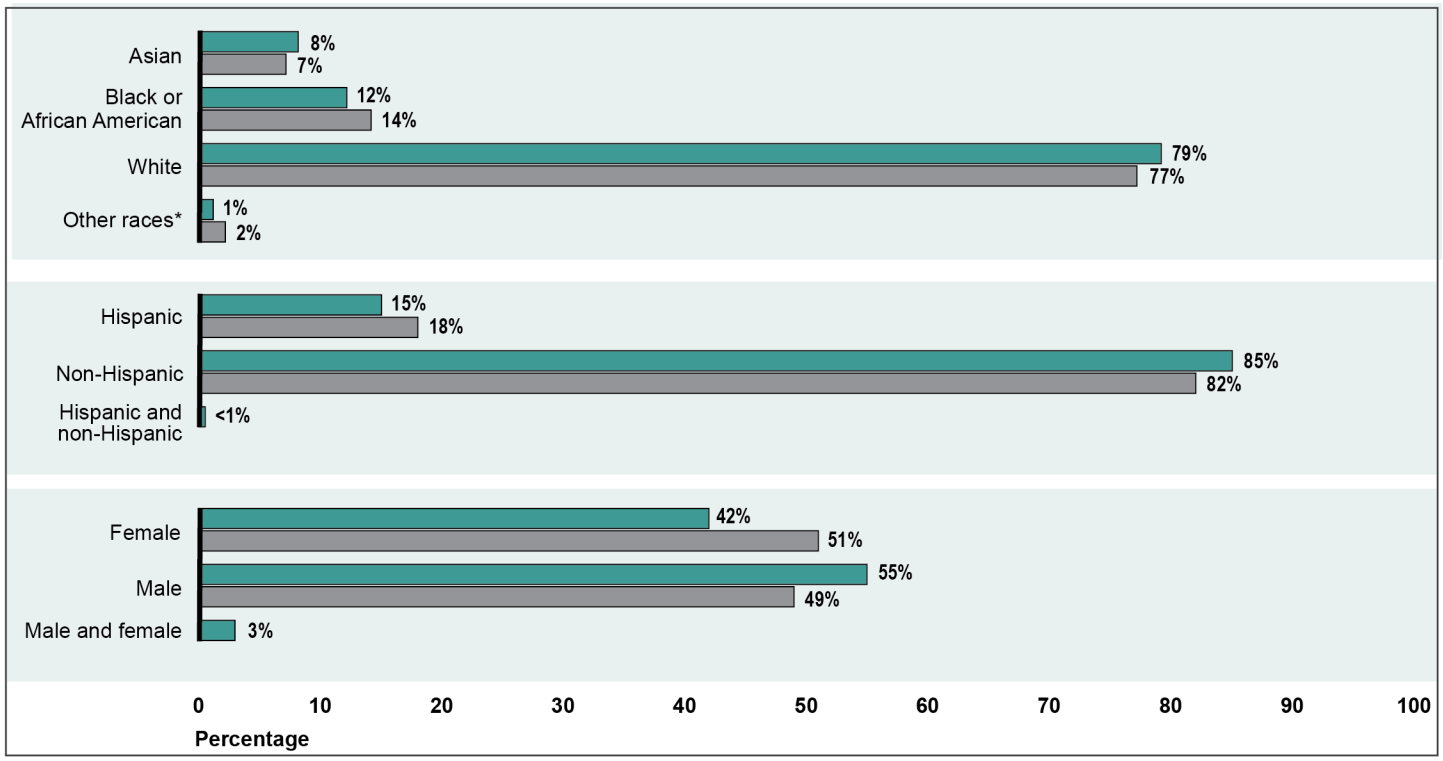
- Black or African American individuals make up approximately 14 percent of the U.S. population. A slightly lower percentage of nonemployer businesses are Black or African American-owned (12 percent). In contrast, only 2 percent of employer businesses are Black or African American owned.[29]
- Hispanic individuals make up approximately 18 percent of the U.S. population. Fifteen percent of nonemployer businesses are Hispanic-owned, while 6 percent of employer businesses are Hispanic owned.
- Female individuals make up about 51 percent of the U.S. population, but 42 percent of nonemployer businesses and 22 percent of employer businesses are female owned.

[27]Demographic data on small business owners come from two sources: Census' 2020 Annual Business Survey (ABS) and 2018 Nonemployer Statistics by Demographics Series (NES-D). The 2020 ABS data survey uses 2019 as the reference year. See our second objective for a discussion of federal sources of demographic data on business owners. Business ownership is defined as having 51 percent or more of the stock or equity in the business. See appendix I for information on how Census categorizes businesses by race, ethnicity, and sex, and how we used the data in our analysis.

[28]A nonemployer business is one without employees and payroll. An employer small business is one with employees and payroll (that is, paid employees). About 81 percent of all small businesses today are nonemployers, according to the SBA Office of Advocacy.

[29]All differences presented in this section (i.e., differences between ABS and American Community Survey (ACS) estimates and between NES-D and ACS estimates) are statistically significant. For the ABS and ACS data analysis, we used reported standard errors to calculate 95 percent confidence intervals around estimates. We determined that differences between estimates were statistically significant if the 95 percent confidence interval surrounding one estimate did not overlap with the 95 percent confidence interval surrounding the other estimate. For NES-D, Census rounds the number of businesses and introduces "noise" or distortions into the data on sales amounts as the primary method of disclosure avoidance. Based on Census' information on rounding and noise flags, we calculated an interval for each estimate that contains the population value. We consider differences between estimates to be significant if the interval surrounding one estimate did not overlap with the interval surrounding the other estimate.

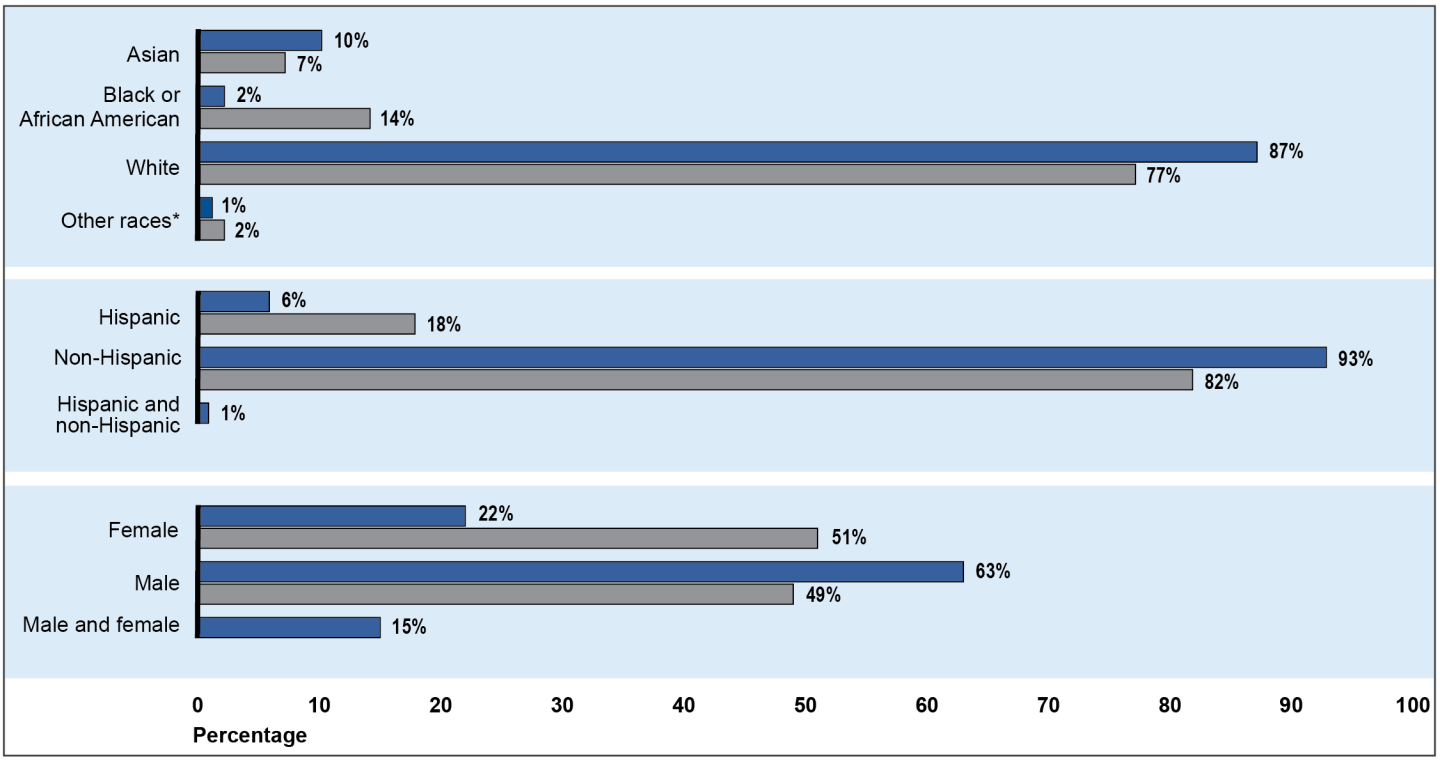
Figure 1: Nonemployer Businesses and the U.S Population by Race, Ethnicity, and Sex



Source: GAO analysis of U.S. Census Bureau data. | GAO-22-104582

Note: A nonemployer business is one without employees and payroll. All national-level demographic estimates had a margin of error of ±0.12 percent or less at the 95 percent confidence level. Individuals who identify as Hispanic may report as any race.

Figure 2: Employer Businesses and the U.S. Population by Race, Ethnicity, and Sex



Source: GAO analysis of U.S. Census Bureau data. | GAO-22-104582

Note: An employer small business is one with employees and payroll (that is, paid employees). All national-level demographic estimates had a margin of error of ±0.12 percent or less at the 95 percent confidence level. All employer demographic estimates had a margin of error of ±0.25 percent or less at the 95 percent confidence level. Individuals who identify as Hispanic may report as any race.

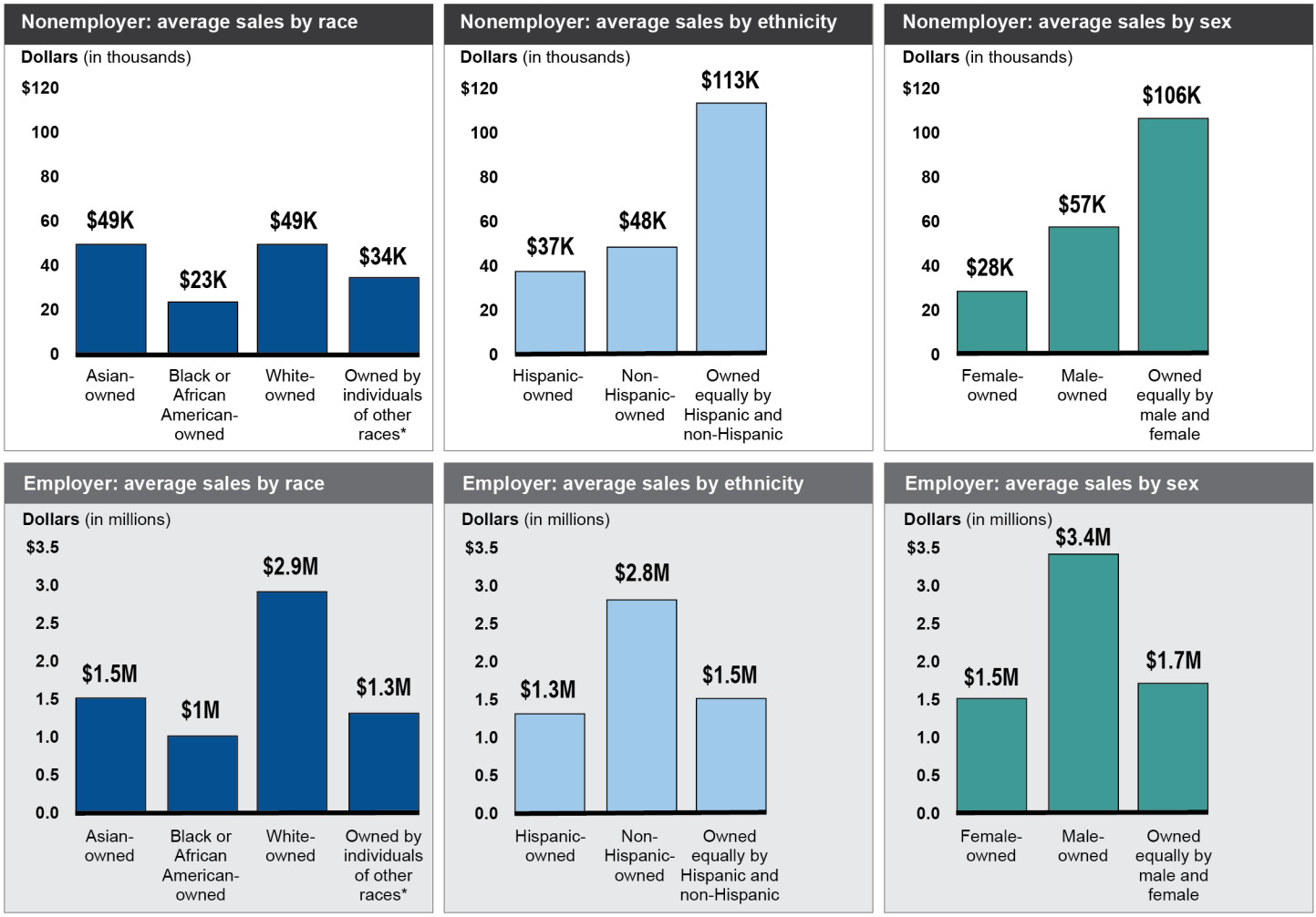
Ownership Rates by Annual Sales

On average, small businesses with Black or African American, Hispanic, or female owners have fewer estimated annual sales than those owned by other demographic groups, according to our analysis of Census data.

- Black or African American-owned nonemployer businesses have about half of the average sales of White-owned nonemployer businesses. Black or African American-owned employer businesses have a third of the average sales of White-owned employer businesses.

- Non-Hispanic-owned nonemployer businesses have about $11,000 more annual sales, on average, than Hispanic-owned nonemployer businesses. Non-Hispanic-owned employer businesses have about $1.5 million more annual sales, on average, than Hispanic-owned employer businesses.
- On average, male-owned businesses have at least double the amount of annual sales compared to female-owned businesses among both employer and nonemployer businesses.
- Nonemployer businesses owned equally by Hispanic and non-Hispanic individuals had much higher average sales than businesses owned by either Hispanic or non-Hispanic individuals. Similarly, nonemployer businesses owned equally by female and male individuals had much higher average annual sales compared to either female or male-owned businesses (see fig. 3). According to Census officials, this difference is likely due to the organizational form of these businesses. Businesses equally owned by two individuals are more likely structured as partnerships or S corporations, as opposed to sole proprietorships, and tend to have higher sales.

Figure 3: Average Sales by Race, Ethnicity, and Sex for Nonemployer and Employer Businesses



*Includes American Indian and Alaska Native and Native Hawaiian and Other Pacific Islander

Source: GAO analysis of U.S. Census Bureau data. | GAO-22-104582

Note: A nonemployer business is one without employees and payroll. An employer small business is one with employees and payroll (that is, paid employees). Nonemployer data are as of 2018; employer data are as of 2019. All employer receipts estimates have a margin of error of ±$152,000 or less at the 95 percent confidence level, with two exceptions. For employer businesses owned by individuals of other races, the margin of error for receipts was ±$525,000. For those owned equally by Hispanic and non-Hispanic individuals, the margin of error was ±$533,000. Both of these figures are at the 95 percent confidence level. Individuals who identify as Hispanic may report as any race.

Small businesses with White, non-Hispanic, or male owners all have an estimated greater share of sales than those with owners of other demographic groups, compared to their representation in the business

owner population. For example, 87 percent of employer businesses are White-owned, and these businesses earn 93 percent of annual sales.

Ownership Rates by Sectors Hit Hardest by the COVID-19 Pandemic

We previously found that businesses in the following six sectors were most likely to experience adverse effects to their business operations as a result of the COVID-19 pandemic:[30]

- accommodation and food services;
- arts, entertainment, and recreation;
- educational services;
- health care;
- manufacturing; and
- retail trade.

Our 2021 analysis used data from the U.S. Bureau of Labor Statistics' 2020 Business Response Survey. Adverse effects to business operations could include a shortage of supplies or inputs, decreased demand for products or services, difficulty moving or shipping goods, and government-mandated closures of business locations.

Our analysis of Census data on business ownership found that certain demographic groups had higher relative rates of small business ownership in sectors hardest-hit by the COVID-19 pandemic, as compared to the demographic makeup of business ownership across all sectors (see fig. 4).[31] This result was found among both nonemployer and employer businesses, although the groups affected varied between these two types of small businesses (see table 12 in appendix III for additional detail). For example:

- Among nonemployer businesses, several demographic groups were overrepresented in four of the six hardest-hit sectors, compared to their representation across all sectors. These groups included White-owned businesses, non-Hispanic-owned businesses, and female-owned businesses. For example, female-owned businesses represented an estimated 76 percent of nonemployer businesses in

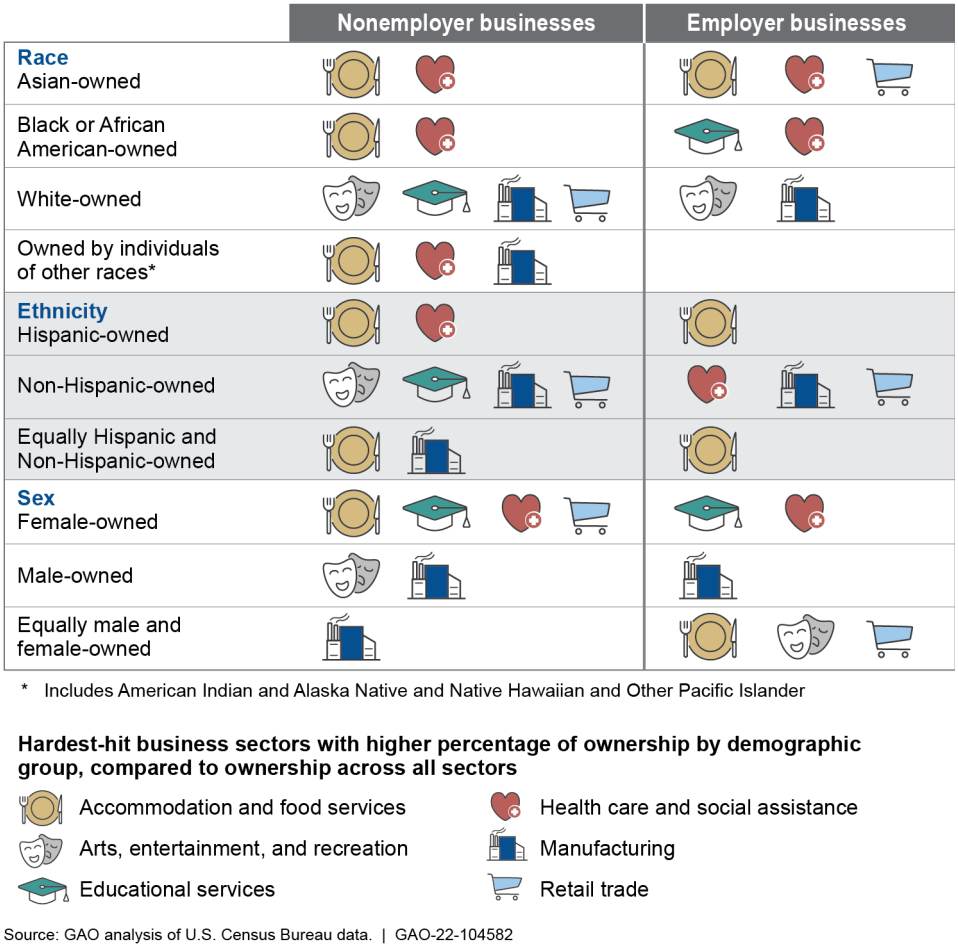
[30]GAO, *Paycheck Protection Program: Program Changes Increased Lending to the Smallest Businesses and in Underserved Locations*, GAO-21-601 (Washington, D.C.: Sept. 21, 2021).

[31]Demographic data on business ownership by sector are from Census' 2020 ABS and 2018 NES-D. The 2020 ABS data survey uses 2019 as the reference year.

the health care and social assistance sector, compared to the percentage of female-owned businesses across all sectors (42 percent).

- Similarly, among employer businesses, demographic groups that were overrepresented in three of the six hardest-hit sectors included Asian-owned businesses, non-Hispanic-owned businesses, and businesses owned equally by males and females. For example, Asian-owned businesses represented an estimated 26 percent of businesses in the accommodation and food services sector, compared to the percentage of Asian-owned businesses across all sectors (about 10 percent).

Figure 4: Higher Business Ownership Rates by Race, Ethnicity, and Sex in Six Sectors Most Adversely Affected by the COVID-19 Pandemic



Source: GAO analysis of U.S. Census Bureau data. | GAO-22-104582

Note: A nonemployer business is one without employees and payroll. An employer small business is one with employees and payroll (that is, paid employees). Nonemployer data are as of 2018; employer data are as of 2019. Sector-level ownership rates by demographic group reported as higher are statistically significant, compared to ownership across all sectors at the 95 percent confidence interval for employer businesses, and based on calculated intervals for nonemployer businesses. Businesses in these six sectors were most likely to experience adverse effects to their business operations as a result of the COVID-19 pandemic, according to the U.S. Bureau of Labor Statistics' 2020 Business Response Survey. Some sector-level data on employer businesses are suppressed. Thus, they are unavailable to report. This affects one sector owned equally by Hispanic and non-Hispanic individuals and two sectors owned by individuals of other races. No ethnicity group had a statistically higher representation among employer businesses in the educational services and arts, entertainment, and recreation sectors. Estimates for non-Hispanic businesses in these two sectors had relatively large margins of error. Individuals who identify as Hispanic may report as any race.

Agency Information and Analytical Methods Can Help Identify Small Business Owners' Race, Ethnicity, and Sex for Tax Analysis Purposes

Federal Agencies' Demographic Data Can Help Identify Disparities in Small Business Owners' Tax Provision Use

The IRS does not collect race, ethnicity, or sex data on taxpayers, including small business owners. However, other federal agencies collect data that contain such demographic information. Particular analytical methods, described later in this report, can help identify or estimate race, ethnicity, and sex information for taxpayers using identified datasets from other federal agencies. These methods could then be used to identify and monitor disparities by race, ethnicity, and sex in the use of tax provisions when applied to federal data.

We interviewed agency officials and reviewed agency websites, technical documentation, and relevant research. We did this to identify relevant datasets from principal federal statistical agencies and agencies with missions that include collecting data on small businesses.[32] Data were

[32]The Office of Management and Budget identified 13 federal statistical agencies. These include the U.S. Census Bureau, the Statistics of Income division at IRS, and the Bureau of Labor Statistics, among others. Federal agencies with missions that include collecting data on small businesses include the Small Business Administration and the Federal Reserve System.

determined to be relevant for our purposes if they included information on business ownership as well as race, ethnicity, or sex information. Some data from the sources we identified, such as the Annual Business Survey and Small Business Credit survey, are collected for statistical purposes, and use results from a sample to generalize to a broad population. In another example, the Department of Agriculture conducts a census of farmers and ranchers every 5 years, including demographic information on these business owners. See figure 5 below for further details on data sources from selected agencies.

Figure 5: Selected Federal Data Sources that Include Race, Ethnicity, and Sex of Business Owners

Data name	Agency	Data collection method	Frequency of collection	Population included	Examples of data on other business characteristics or economic conditions
Annual Business Survey	U.S. Census Bureau	Survey	Annual	Employer businesses with more than $1,000 in sales	• Sales • Number of employees • Industry • Geographic location
Nonemployer Statistics by Demographics	U.S. Census Bureau	Administrative	Annual	Businesses with no paid employment or payroll and annual sales of at least $1,000, or $1 in construction industries	• Sales • Legal form of business • Industry • Geographic location
Economic Census of Island Areas	U.S. Census Bureau	Census	Every 5 years	Businesses in American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, the U.S. Virgin Islands, and Puerto Rico	• Operating revenue • Number of employees • Industry • Legal form of business
Small Business Credit Survey	Federal Reserve Banks	Survey	Annual	Businesses with fewer than 500 employees	• Revenue • Number of employees • Industry • COVID-19 assistance received
Census of Agriculture	Department of Agriculture	Census	Every 5 years	All farms with more than $1,000 of sales or potential sales of agricultural products	• Income and sales • Legal form of business
Agricultural Resource Management Survey	Department of Agriculture	Survey	Annual	All farms with more than $1,000 of sales or potential sales of agricultural products	• Income and sales • COVID-19 assistance received • Legal form of business

Source: GAO analysis of information from the Department of Agriculture, Federal Reserve Banks, and U.S. Census Bureau. | GAO-22-104582

Additionally, agency administrative data may contain race, ethnicity, or sex information captured on benefit applications or collected during use of services. For example, the Small Business Administration (SBA) collects data from lending institutions that distribute SBA loans. These data

include aggregate loan amounts by race, ethnicity, and sex of loan recipients. Providing this demographic data is optional, and does not include the race, ethnicity, and sex of every loan recipient.

Federal data sources that offer important economic information on current business conditions can be analyzed alongside demographic data sources to provide information. For example, the Small Business Pulse Survey is a federal statistical survey conducted by Census to measure small business experiences during the COVID-19 pandemic. In our prior work, we used data from this survey to describe variation by industry in the number of businesses that reported receiving certain COVID-19 tax provisions.[33] For our March 2021 report, we also analyzed race, ethnicity, and sex characteristics of business owners by industry, using data from the Annual Business Survey (ABS). We were unable to draw any conclusions about the race, ethnicity, or sex of the business owners that reported receiving the COVID-19 tax provisions because we did not have business-level demographic data for owners of businesses receiving the credits.

There are challenges in collecting and analyzing information on the race, ethnicity, and sex of business owners in existing government datasets, particularly when assigning demographic characteristics to multi-owner businesses. For example, ABS defines ownership as having 51 percent or more of the stock or equity in the business. The Census of Agriculture counts demographic characteristics of all producers, which are defined as any individuals who were involved in making decisions for the enterprise. The U.S. Census Bureau assigns businesses to a demographic group if owners of that group collectively own a majority stake in the business.[34]

Although not specific to measuring the demographics of business owners, voluntary survey compliance presents an additional challenge. If the demographic characteristics of respondents and nonrespondents vary systematically, the measured demographics among respondents could be similarly biased. Some selected agencies attempt to avoid bias due to

[33]GAO, *COVID-19: Sustained Federal Action Is Crucial as Pandemic Enters Its Second Year*, GAO-21-387 (Washington, D.C.: Mar. 31, 2021).

[34]Businesses cannot be classified by demographics in some instances, such as when no one group collectively owns 10 percent or more of the business or if it is owned by another business.

voluntary survey compliance by using imputation to address gaps in survey data on race, ethnicity, and sex.

Demographic Data Can Be Identified by a Variety of Methods

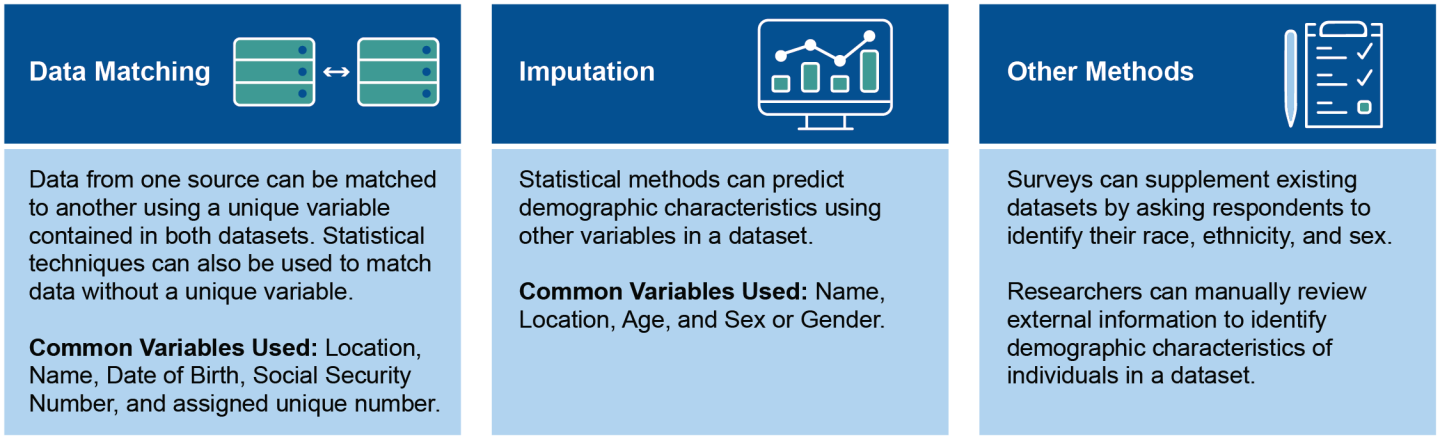
Multiple analytical methods used in academic research can help identify race, ethnicity, and sex information when it is missing from a dataset. Applying these methods to relevant government data sets can estimate demographic characteristics for taxpayers. This, in turn, can help determine whether there are disparities in the use of tax benefits by demographic group. We previously reported that nontax data suggest the potential for disparities in the use of wealth-oriented tax provisions.[35] Analysis of demographic data could help identify whether there is differential use of tax provisions among small business owners from diverse backgrounds. These data can also help determine the extent to which differential use may exacerbate existing inequities in business outcomes.

We reviewed studies to identify methods researchers have used that could be applicable to identifying taxpayer race, ethnicity, or sex.[36] Specifically, we reviewed 39 studies from peer-reviewed journals that were published within the last 10 years and used a separate data source to assign missing race, ethnicity, and sex information to administrative data. We identified methods used in these studies, as well as their strengths and limitations. We also provide a nonexhaustive list of examples to illustrate the types of challenges faced when using various methods.

[35]GAO, *Tax Equity: Lack of Data Limits Ability to Analyze Effects of Tax Policies on Households by Demographic Characteristics,* GAO-22-104553 (Washington, D.C.: May 18, 2022).

[36]We use the term “sex” to describe the missing demographic data for studies that used analytical methods to obtain missing sex or gender data.

Figure 6: Examples of Methods to Identify Missing Race, Ethnicity, and Sex Data



Source: GAO analysis of academic literature. | GAO-22-104582

Studies we reviewed used methods to identify potential disparities in areas where no demographic data had been collected. Out of the 39 studies we reviewed, 31 had research objectives that directly related to examining inequalities or disparities by race, ethnicity, or sex. For example, one study matched income tax data with Census information to identify income inequality and mobility across racial and ethnic groups.[37] Another study validated an imputation method against self-reported race and ethnicity information to improve research into occupational health disparities among workers from different backgrounds.[38]

Data Matching

Out of 39 studies we reviewed, 17 combined datasets to identify race, ethnicity, or sex information for individuals. In these studies, relevant demographic data was missing in one dataset and present in the other. Most of these studies matched datasets using a unique identifier present in both datasets when available, such as location, name, date of birth, or Social Security number. Seven of the studies used statistical techniques to match data without a unique identifier. These studies used various factors such as location, name, age, or sex to predict the probability that individual records from different sources matched. Five studies also

[37]Randall Akee, Maggie R. Jones, Sonya R. Porter, "Race Matters: Income Shares, Income Inequality, and Income Mobility for All U.S. Races," *Demography,* vol. 56, no. 3 (2019): 999.

[38]Caroline K Smith, David K. Bonauto, "Improving Occupational Health Disparity Research: Testing a Method to Estimate Race and Ethnicity in a Working Population," *American Journal of Industrial Medicine* vol. 61, no. 8 (August 2018): 641, 646.

combined these methods, matching some data using a unique identifier, and then using probabilistic techniques to match the remaining records.

Six of the matching studies we reviewed used federal datasets, such as those developed by Census or the Social Security Administration (SSA), to identify the race, ethnicity, and sex of individuals. Examples of federal data sources that provide demographic information include Census' American Community and Current Population Surveys, and Decennial Census. Other examples of data sources include the Survey of Consumer Finances from the Board of Governors of the Federal Reserve System and administrative data collected by SSA.[39] Studies we reviewed were able to use nonpublic federal data through methods such as using data matched by the agency or assigned a unique anonymized identifier prior to sharing.

There are challenges in matching data sources, according to the studies we reviewed and other research.[40] The National Academy of Sciences identified that matching methodologies should include considerations of completeness, accuracy, timeliness, and purpose. Studies we reviewed also described inconsistencies in individuals' race and ethnicity identification across datasets. For example, one study found that individuals under 18 years old were more likely to have missing race and Hispanic origin data in administrative records data, when compared to the 2010 Census results.[41]

In cases where matches between two records could not be determined, studies generally excluded the records from the analysis, potentially affecting the completeness of the information available for particular groups. One study supplemented its analysis with datasets that had high

[39]SSA collects data on individuals' sex. According to an author from SSA's Office of Retirement and Disability Policy, there are reliability concerns with SSA's race and ethnicity data. Administrative data, as defined by the Office of Management and Budget, are administrative, regulatory, law enforcement, adjudicatory, financial, or other data typically collected and held by agencies to carry out the basic administration of a program, such as processing benefit applications or tracking services received.

[40]National Academies of Sciences, Engineering, and Medicine, *Innovations in Federal Statistics: Combining Data Sources While Protecting Privacy* (Washington, D.C.: The National Academies Press, 2017).

[41]Sharon R. Ennis, Sonya R. Porter, James M. Noon, Ellen Zapata, "When Race and Hispanic Origin Reporting are Discrepant across Administrative Records and Third Party Sources: Exploring Methods to Assign Responses," *Statistical Journal of the IAOS*, vol. 34, no. 2 (2018): 187.

agreement for Hispanic identification to mitigate this limitation.[42] Additionally, Treasury officials told us that matching data from existing surveys can be a challenge because surveys are less complete than tax data.

Imputation

We also identified imputation as a method of estimating race, ethnicity, and sex. Imputation uses statistical models to assign values based on information available in the dataset. Out of the 39 studies we reviewed, 29 used some form of imputation to estimate race, ethnicity, or sex. Studies used imputation to estimate race, ethnicity, and sex data for various groups, including medical patients, voters, and business owners.

These studies used variables including residential location, name, age, or industry to develop their estimates. For example, one study estimated the ethnicity of cancer patients in the United Kingdom using name recognition software and location information.[43] Another study used the name and location variables combined with the age, primary language, and self-reported race of family members to estimate the race of medical patients.[44]

Studies used different imputation methodologies including approaches such as statistical modeling, identifying patterns from large datasets, combining probabilities from multiple variables, or transforming probabilities into an assigned category. One imputation method known as Bayesian Improved Surname Geocoding (BISG) predicts race and ethnicity by calculating the probability that a person with a given surname and residential location will identify with selected racial and ethnic groups, based on data from Census. Of the 29 studies that used imputation, 14 used BISG or a variation on the method. For example, one study

[42]Ennis, Porter, Noon, and Zapata, "When Race and Hispanic Origin Reporting are Discrepant," 181.

[43]Ronan Ryan, Sally Vernon, Gill Lawrence, Sue Wilson, "Use of Name Recognition Software, Census Data and Multiple Imputation to Predict Missing Data on Ethnicity: Application to Cancer Registry Records," *BMC Medical Informatics and Decision Making*, vol.12, no. 3 (2012): 1.

[44]Gabrielle C. Silva, Amal N. Trivedi, Roee Gutman, "Developing and Evaluating Methods to Impute Race/Ethnicity in an Incomplete Dataset," *Health Services and Outcomes Research Methodology*, vol.19 (2019): 175, 177, 178.

incorporated first names into its analysis in addition to the surnames typically used for BISG.[45]

Imputation studies overall most commonly used name, location, or age to predict race, ethnicity, or sex. However, other predictors, such as medical conditions or industry, were also used. Studies generally used multiple predictors to estimate race and ethnicity, but some studies used one. For example, four studies used name (first, last, or both) as the only predictor to estimate sex. Three others used it as the only predictor to estimate race or ethnicity. Nine studies used age as a factor in estimating race and ethnicity by name and location.

Imputation methods are generally effective in estimating particular race categories, ethnicity, and sex, according to studies we reviewed. We reviewed 21 studies that compared the accuracy of imputation approaches against self-reported data to test the accuracy of the predictions. These results showed high accuracy in general for sex and the largest race categories in the United States (Asian, Black or African American, and White).

While the accuracy of imputation was generally high, it varied across racial and ethnic groups. For example, one study found that its algorithm was most accurate in estimating the race and ethnicity for individuals who were Black or African American (92 percent) or Hispanic (83 percent), but less accurate for those who were Asian or White.[46] Another study provided estimates that were most accurate for those identified as Black or African American (94 percent), Asian (89 percent), and White (86 percent), and slightly less accurate (79 percent) for those who identified as Hispanic.[47] According to one study we reviewed, the predictors and methods used can influence accuracy. The study compared different

[45]Ioan Voicu, "Using First Name Information to Improve Race and Ethnicity Classification," *Statistics and Public Policy,* vol.5, no.1 (2018): 1.

[46]Katie Labgold, Sarah Hamid, Sarita Shah, Neel R. Gandhi, Allison Chamberlain, Fazle Khan, Shamimul Khan, Sasha Smith, Steve Williams, Timothy L. Lash, and Lindsay J. Collin, "Estimating the Unknown: Greater Racial and Ethnic Disparities in COVID-19 Burden After Accounting for Missing Race and Ethnicity Data," *Epidemiology*, vol.32, no. 2 (2021): 159.

[47]Anna Haas, Marc N. Elliott, Jacob W. Dembosky, John L. Adams, Shondelle M. Wilson-Frederick, Joshua S. Mallett, Sarah Gaillot, Samuel C. Haffer, Amelia M. Haviland, "Imputation of Race/Ethnicity to Enable Measurement of HEDIS Performance by Race/Ethnicity," *Health Services Research,* vol. 54 (2019): 22.

methods and found that imputation models that used family members' race as a predictor performed better.[48] Multiple studies also showed that imputation methods had low accuracy in estimating race for groups with smaller populations, such as American Indian and Alaskan Native, and those identifying as more than one race.

There are general limitations with using imputation methods to develop taxpayer data that include demographic information. Imputation methods can produce estimates with some error. This could affect the reliability of summary statistics on tax outcomes by race, ethnicity, and sex. Imputation methods can also introduce bias into the data. This could produce inaccurate conclusions about the correlations between demographic factors and tax outcomes. For example, if analysts used income to impute race and ethnicity, then correlations among race and tax outcomes may actually reflect correlations with income, the variable used for imputation. These limitations might become more pronounced when imputations are used to conduct detailed analyses of specific tax provisions.

Other Methods

We identified a small number of studies in our review that used other methods to identify race, ethnicity, and sex information when not included in a dataset. One study we reviewed described using surveys to obtain new race and ethnicity information or supplement existing information. We described surveying taxpayers as an option for IRS to collect demographic information in our May 2022 report on tax equity.[49] We also identified two studies that manually classified information to estimate data on sex. These studies used human coders to review names and identify sex through web research. Manual classification would require significant resources to use for larger datasets.

Original data collection is another way to identify the race, ethnicity, and sex of taxpayers, but there are challenges to this approach. In our May 2022 tax equity report, we described adding demographic questions to tax forms as an option that IRS could use to identify the race, ethnicity, and sex of taxpayers.[50] We also interviewed IRS officials who said that any direct collection of demographic information by the agency could significantly compromise voluntary compliance. Experts we interviewed

[48]Gabrielle C. Silva, Amal N. Trivedi, Roee Gutman, "Methods to Impute Race/Ethnicity," 175.

[49]GAO-22-104553.

[50]GAO-22-104553.

for the report also cited concerns with public reaction and the potential for inadvertent consequences of IRS examiners having access to that information. However, some experts acknowledged that there might be ways of safeguarding that information so it is used only for research purposes.

Federal Agencies Have Engaged in Some Efforts to Analyze Tax Data by Demographic Categories

While there is no source of data that consistently connects federal tax information to the race, ethnicity, and sex of the taxpayer, Treasury and IRS have analyzed the use of some tax provisions by demographic characteristics. For example, Census and IRS have collaborated in the past to match taxpayer and survey data. In 2021, IRS presented a paper in which it used 2010 tax data matched with Census data to estimate the extent of individual income tax nonfiling.[51]

There are limitations on sharing taxpayer data, including legal protections that require the information remain confidential. For example, Titles 13 and 26 of the United States Code limit the ability of Census and IRS to share data, respectively. Although Title 13 permits Census to enter into statistical project agreements, it also limits the use of the information Census has collected to the statistical purposes for which it was supplied. Title 26 contains broad protections for taxpayer data.[52] To meet their legal requirements, Census and IRS enter into specific project-based sharing agreements that, according to officials from those agencies, require detailed, often legal, reviews that can involve significant resources and time. These projects must also maintain the confidentiality of data.[53] IRS

[51]Tom Hertz, Pat Langetieg, Mark Payne, and Alan Plumley, *New Approaches to Estimating the Extent of Nonfiling* (paper presented at the 11th Annual Internal Revenue Service-Tax Policy Center Joint Research Conference on Tax Administration, June 24, 2021).

[52]26 U.S.C. § 6103 provides that all returns and return information shall be confidential and shall not be disclosed, subject to limited exceptions listed in the section or authorized elsewhere under Title 26 of the United States Code. One such exception is section 6103(j)(1) of Title 26. It requires IRS to share federal tax information with Census for the statistical purposes of, but only to the extent necessary in, the structuring of censuses and national economic accounts, as well as for conducting related statistical activities authorized by law. This section of the code has a related regulation found at 26 C.F.R. § 301.6103(j)(1)-1.

[53]These projects between Census and IRS have used a process designed to protect individuals' identities with no direct link to tax data. That process assigns a unique identifier to Census and IRS data and uses it to match the files without the use of private information.

and Treasury officials agreed that continued work to produce tax analyses in connection with demographic information is important.

According to IRS officials, Census and IRS have had an interagency agreement since 2005 to produce annual estimates of Earned Income Tax Credit participation. According to Census officials, Census uses IRS administrative tax records and Census survey data to estimate characteristics of eligible and participating taxpayers, including race, ethnicity, and sex. IRS officials also said that they are currently working with Census to establish a new interagency agreement to produce information, including taxpayer racial characteristics, on eligibility and use of the Child Tax Credit. The current estimate for delivery is 2023, according to IRS officials. Additionally, Treasury officials told us that Treasury has matched taxpayer and SSA data to analyze taxpayers' use of health insurance through public exchanges by sex.

Treasury has not generally analyzed taxpayer data by demographic categories such as race and ethnicity, according to officials, but has considered options to begin this type of analysis. Recent guidance from President Biden and the Office of Management and Budget has directed agencies to consider issues of equity for their policies and programs by using demographic data collection and analysis.[54] In December 2021, Treasury announced that it is analyzing when and how adults and families accessed economic impact payments (EIP), including analysis by demographic categories.[55] Specifically, Treasury plans to analyze how outreach affected equity in accessing the first-round EIP.

Treasury is also working to develop a reliable imputation method to estimate taxpayer race and ethnicity in general to evaluate the effectiveness and equity of tax provisions. IRS officials told us that having demographic data could be useful for administering the tax code by helping to inform outreach strategies aimed at taxpayers of different demographic groups. Our recent report on tax equity recommended that Treasury evaluate the feasibility of alternative methods, such as inter-

[54]Office of Management and Budget, *Evidence-Based Policymaking: Learning Agendas and Annual Evaluation Plans*, OMB Memorandum M-21-27 (Washington, D.C.: June 30, 2021); White House, *Memorandum on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking* (Washington, D.C.: Jan. 27, 2021); and Exec. Order No. 13985, 86 Fed. Reg. 7009 (Jan. 25, 2021).

[55]U.S. Department of the Treasury, *Advancing Equity Analysis in Tax Policy,* https://home.treasury.gov/news/featured-stories/advancing-equity-analysis-in-tax-policy.

agency data sharing or surveys, for producing secure, linked taxpayer and demographic data.[56]

Use of COVID-19 Credits and Deferrals Varied by Race, Ethnicity, and Sex of Small Business Owners Based on Our Analysis of a Study Population

To demonstrate the application and feasibility of tax-related differential use analyses, we used two methods to obtain missing demographic data for taxpayers. We then estimated differences in the use of selected COVID-19 tax provisions by small business owners' race, ethnicity, and sex. We used a data matching method to identify the sex of business owners and an imputation method to identify race and ethnicity. Using these methods, we provide estimates on the use of COVID-19 tax provisions among a selected study population of businesses. However, the results are not generalizable to all businesses filing tax returns. We also identified limitations to consider when conducting similar disparity analyses.

Feasibility and Methodological Considerations for Analysis of Small Business Tax Provision Use

To determine feasibility of a differential use analysis, we selected provisions to examine, identified a study population, and estimated provision eligibility. Our analysis examined the use of three COVID-19 tax provisions for employers: the Employee Retention Credit (ERC), paid sick and family leave credits (leave credits), and the deferrals of tax payments for the employer's share of payroll tax.[57] Self-employed business owners were also eligible for the leave credits and payroll tax deferral. We provide separate results for employers and self-employed business owners for both provisions.[58] Our analysis covers use of these provisions

[56]GAO-22-104553. Treasury neither agreed nor disagreed with the recommendation, saying it had considered other options. We recognize Treasury's efforts but maintain that implementing our recommendation would better position Treasury to establish linked taxpayer and demographic data that could be used to analyze the effects of tax policies by race, ethnicity, and sex.

[57]On August 8, 2020, a Presidential Memorandum directed the Secretary of the Treasury to use his authority to allow the deferral of withholding, deposit, and payment of the employee share of certain employment taxes imposed on wages or compensation paid from September 1, 2020, through December 31, 2020. In May 2022, we found that 99 percent of the amount deferred by employers in 2020 was for the employer share of payroll tax. See GAO-22-104280.

[58]In May 2022, we found that while self-employed individuals made up the majority of users of the leave credits and payroll tax deferral, the credit and deferral amounts were much smaller for self-employed individuals, as compared to employers. See GAO-22-104280.

on 2020 annual returns and on quarterly returns for the second through fourth quarters of 2020.

Our study population did not include all users of the selected COVID-19 tax provisions. In addition, because we could not select users with known probabilities, our analysis did not produce results that generalize to the universe of businesses using the provisions. Assigning demographics to businesses with multiple owners is difficult and requires complex decision rules on determining shares of ownership and assigning demographics to the business unit accordingly.

For ease of analysis in this demonstration, we limited our study population to businesses with a single owner. Specifically, our selected study population of small businesses included (1) sole proprietors filing form 1040, Schedule C or F; and (2) S corporations where a single individual was identified across ownership forms.[59] We excluded partnerships and C corporations because these business forms are more likely to have complex ownership structures and ownership that is difficult to determine based on tax returns alone. In addition, to the extent possible, we excluded sole proprietorships operated as a qualified joint venture by a married couple.[60]

We further narrowed our study population based on the provisions we analyzed. For analysis of the employer provisions, our study population consisted of approximately 2.8 million businesses filing an employment tax return. This represents approximately 39 percent of all employment tax returns in the second through fourth quarters of 2020.[61] For analysis

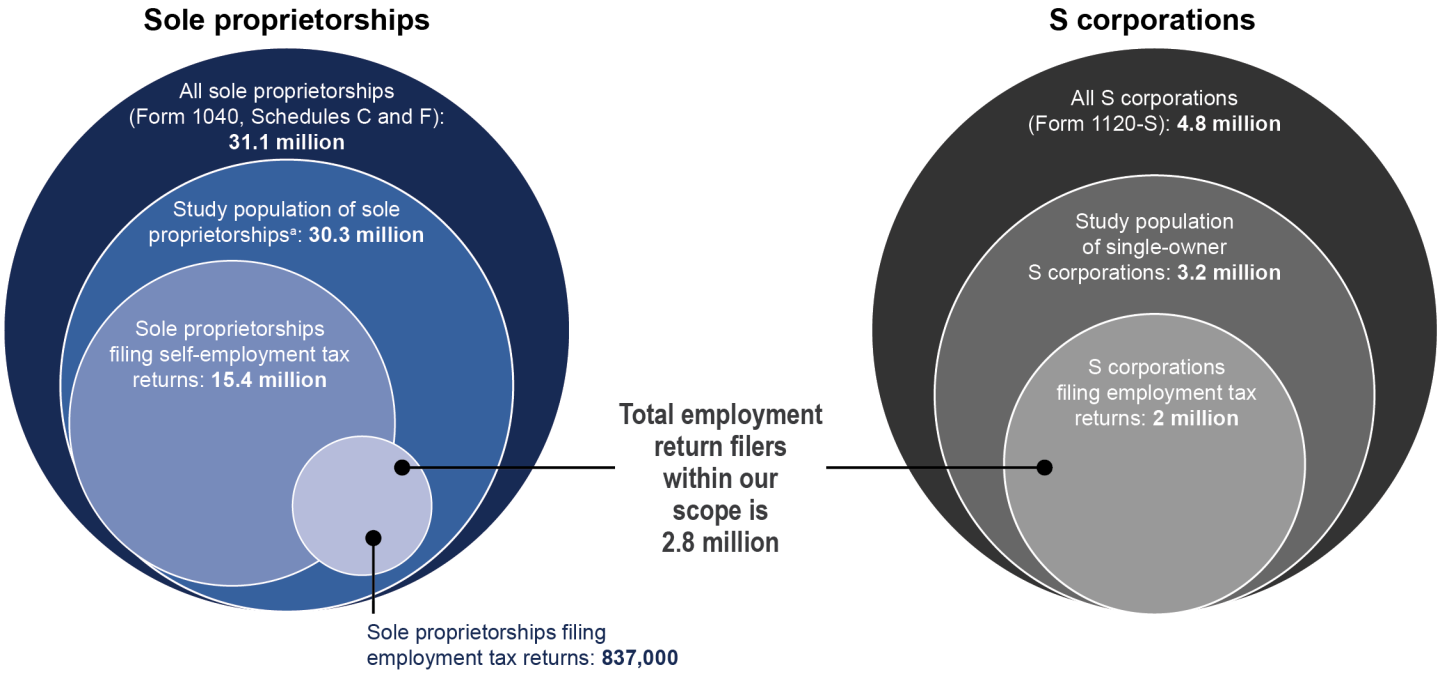
[59]Data on ownership of S corporations were obtained from Form 1120-S, Schedule K-1, *Shareholder's Share of Income, Deductions, Credits, etc.* Corporations use Schedule K-1 to report individuals' shares of the corporation's income, deductions, credits, and other items. We did not include S corporations in our analysis of the self-employed provisions because S corporations pay employment taxes through employment tax returns rather than self-employment tax returns.

[60]According to an IRS website, a qualified joint venture is one that conducts a trade or business where (1) the only members of the joint venture are a married couple who file a joint return, (2) both spouses materially participate in the trade or business, and (3) both spouses elect not to be treated as a partnership. http://www.irs.gov/businesses/small-businesses-self-employed/election-for-married-couples-unincorporated-businesses.

[61]This count of employment tax returns included unique employer identification numbers across Forms 941, 943, and 944. It also included businesses that used Form 941, Schedule R, to file employment tax returns. Data are as of September 2021 to March 2022.

of the self-employed provisions, our study population consisted of approximately 15.4 million sole proprietorships with a Form 1040, Schedule SE, in 2020 (see fig. 7). These businesses are owned by 14.2 million owners, representing 78 percent of all Schedule SE filings in 2020.[62] For further details about our study population, including demographic information, see appendixes I and II.

Figure 7: Small Business Study Population for Use Analysis of COVID-19 Tax Provisions, 2020



Source: GAO analysis of Internal Revenue Service taxpayer data. | GAO-22-104582

Note: Employment tax returns include forms 941, 943, and 944. Self-employment tax returns include form 1040, Schedule SE. Overlap between self-employed sole proprietorships and employer sole proprietorships was approximately 504,000 businesses in 2020.

[62]According to IRS, generally someone is self-employed if they (1) carry on a trade or business as a sole proprietor or an independent contractor; (2) are a member of a partnership that carries on a trade or business; or (3) are otherwise in business for themselves (including a part-time business). Because our focus was to analyze demographics of eligible provision users, we did not apply any tests based on income or deductions to narrow the population to those engaged in business activity, as it is traditionally understood. Our study population is based on tax forms that report business income. The unit of analysis is business level rather than the individual level. Therefore, we refer to the self-employed as "business owners" in discussion of the results. Some self-employed business owners (approximately 7.5 percent of the 14.2 million owners) had more than one business that fell within our study population during tax year 2020. In these cases, the provision use is counted once per business.

[a]Study population for sole proprietorships does not include businesses identified as qualified joint ventures or businesses with nonunique employer identification numbers.

Not all businesses within our study population were eligible for each selected tax provision. To represent use of the provisions accurately by demographic group, we estimated an eligible population of business owners for almost all provisions, using available information within the taxpayer data, and based on IRS guidance for employers and self-employed business owners. We conducted statistical tests to ensure that our eligibility decisions captured a sufficient percentage of businesses within our study population that used the provisions.

We were unable to determine eligibility for the ERC due to eligibility rule complexity and lack of data on quarterly receipts and government-ordered suspension of operations. For the ERC, we compare the demographic makeup of business owners filing employment tax returns within our study population to that of business owners using the ERC.

As discussed previously, demographics of business ownership vary by business size. To control for the effects of business size in our analysis, we examined annual business receipts, as reported on tax returns for 2020, when estimating how usage rates varied across demographic groups.[63]

Use of COVID-19 Tax Provisions by Sex of Small Business Owners

Matching across administrative datasets is one way to address missing demographic data. For our study population of businesses, we matched owner Social Security numbers, as provided on the relevant tax forms, to SSA data to identify the individual's recorded sex.[64] We matched 99.96 percent of business owners within our study population as either female or male. Using the identified sex of business owners, we analyzed

[63]Tax return fields used for receipts analysis vary by form. For Form 1040, Schedule C, and 1120-S, we analyzed gross receipts data from lines 1 and 1a, respectively. For Form 1040, Schedule F, there are multiple lines for reporting various receipts and income sources and not all taxpayers report values in all these fields. Therefore, we analyzed the field that was most populated—the gross income from the cash method, which is reported on line 9. For some businesses filing a Form 1120-S, 2020 receipts data were unavailable.

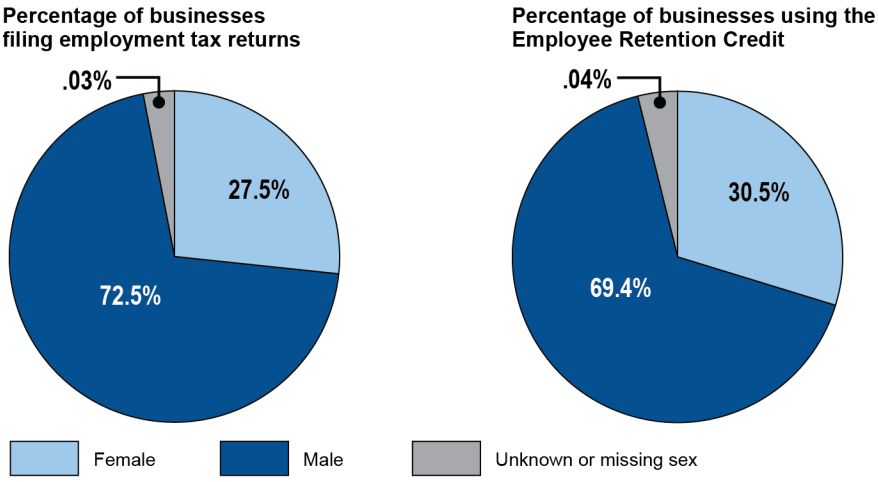
[64]We accessed SSA data on sex through IRS's Compliance Data Warehouse. We analyzed the effect of COVID-19 tax provisions by sex, which, for the purposes of this report, includes the variables female and male. We use the terms female and male because the dataset we analyze uses these terms to describe sex variables. Furthermore, our analysis does not consider gender, as it signifies social and cultural factors absent from our data.

whether differences by sex existed in the use of the selected tax provisions among the study population of businesses.

Employee Retention Credit

Within the study population, female-owned businesses made up a slightly higher percentage of ERC users (30.5 percent) than of businesses filing employment tax returns (27.5 percent) during tax year 2020 (see fig. 8). Male-owned businesses were the majority among ERC users. However, relative to the number of male-owned businesses filing employment tax returns, slightly fewer male-owned businesses used the ERC. This result occurred across all quartiles of business size (see appendix III, table 13).

Figure 8: Businesses Filing Employment Returns and Use of the Employee Retention Credit by Sex, among the Study Population, 2020



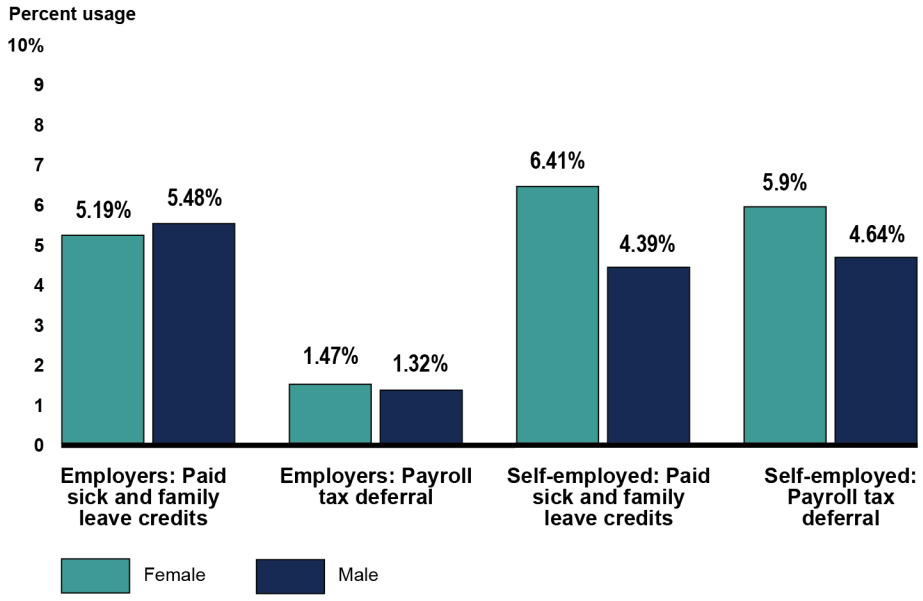
Source: GAO analysis of Internal Revenue Service taxpayer data and Social Security Administration data. | GAO-22-104582

Note: Taxpayer data are as reported by taxpayers and subject to taxpayer reporting errors. Use of the Employee Retention Credit is as reported on employment tax forms 941, 943, and 944, and form 7200 for advance payment of the COVID-19 tax provisions. Our analysis included credits claimed on Schedule R and through amended returns associated with form 941. Quarterly returns (forms 941 and 7200) from second through fourth quarter 2020 and annual returns for 2020, including electronically filed returns and paper filings, are included in the figure. The study population included businesses filing employment tax returns, which were also either (1) sole proprietorships filing form 1040, Schedule C or F; or (2) S corporations with a single owner. Data are as of September 2021 to March 2022. Results are not generalizable to the universe of businesses using the provision. Numbers may not add to 100 because of rounding.

Leave Credits and Payroll Tax Deferrals

Female and male-owned businesses used the employer leave credits and the employer payroll tax deferral at similar rates in 2020 within the study population (see fig. 9).[65]

Figure 9: Use of Selected COVID-19 Employer and Self-Employed Provisions by Sex, among Eligible Businesses in the Study Population, 2020



Source: GAO analysis of Internal Revenue Service taxpayer data and Social Security Administration data. | GAO-22-104582

Note: Taxpayer data are as reported by taxpayers and subject to taxpayer reporting errors. Use of the employer provisions is as reported on employment tax forms 941, 943, and 944, and form 7200 for advance payment of the COVID-19 tax provisions. Our analysis included credits claimed on Schedule R and through amended returns associated with form 941. Use of the self-employed provisions is as reported on form 1040, Schedule 3. Quarterly returns (forms 941 and 7200) from second through fourth quarter 2020 and annual returns for 2020, including electronically filed returns and paper filings, are included in the figure. For the employer provisions, the study population included businesses filing employment tax returns, which were also either (1) sole proprietorships filing form 1040, Schedule C or F; or (2) S corporations with a single owner. For the self-employed provisions, the study population included sole proprietorships filing form 1040, Schedule C or F, and which also

[65]For the employer paid sick and family leave credits, we used employee counts reported on employment tax forms 941 and 943 to estimate eligibility for the provision. Businesses who reported 500 or more employees in tax year 2020 were excluded from the eligible population. For quarterly returns, we excluded businesses if their average number of employees across the second through fourth quarters of tax year 2020 were 500 or more. Within our study population, about 96 percent of female-owned businesses and 95.6 percent of male-owned businesses filing employment tax returns were eligible for the leave credits. The number of employees is not collected on form 944, an annual employment return form for the smallest of employers (i.e., businesses whose annual liability for Social Security, Medicare, and withheld federal income taxes is $1,000 or less). Given the small size businesses filing form 944, we assume that all are eligible for the employer leave credits.

filed a form 1040, Schedule SE. Data are as of September 2021 to March 2022. Results are not generalizable to the universe of businesses using the provisions.

There was limited variation in use of the employer provisions by sex across different business types and sizes. Specifically, use rates for the employer provisions were similar by sex and business type (e.g., sole proprietorships and S corporations), with differences of less than 1 percentage point. Likewise, use of the employer payroll tax deferral by sex was similar across different business sizes. However, use of the employer leave credits by sex varied slightly when examining these differences across business sizes. Female-owned businesses in the two highest-sized quartiles (i.e., the 50th to 74th and 75th to 100th quartiles) used the provision at slightly greater rates than male-owned businesses. These differences were approximately 2 percentage points.

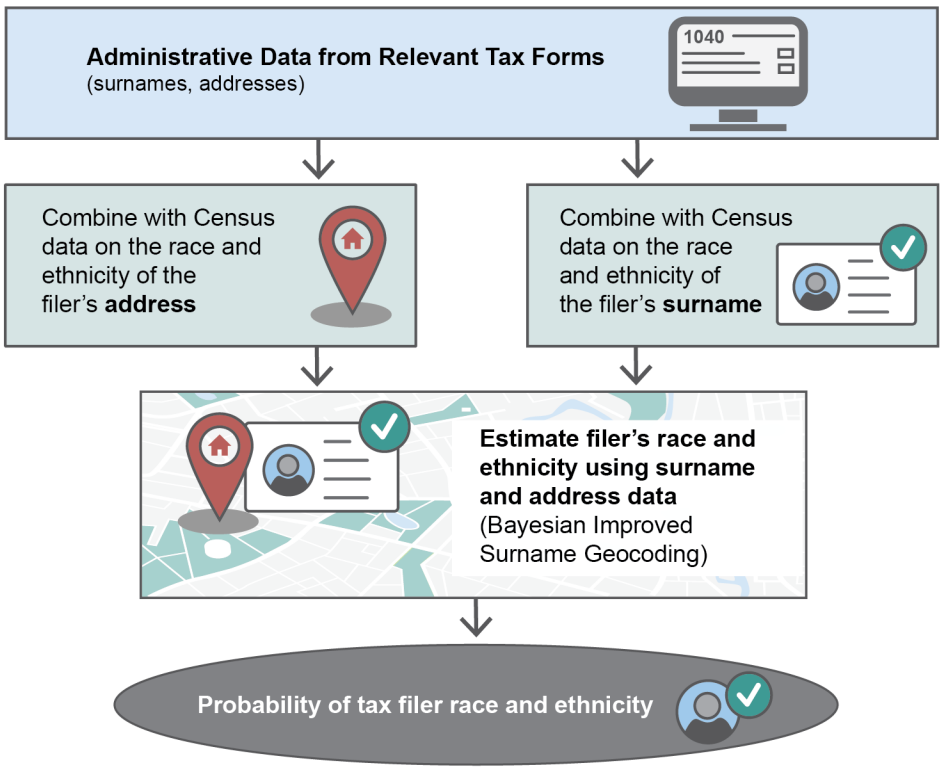
Female-owned businesses used the self-employed leave credits and payroll tax deferral at slightly higher rates than male-owned businesses—a difference of approximately 2 and 1.3 percentage points, respectively. Differences in use of these provisions between female- and male-owned businesses did not vary substantially across business size quartiles. For ratios in provision use by female-to male-owned businesses and by business size quartiles, see appendix III, table 14. Given that our analysis is limited to a subset of businesses, we were unable to determine if differences in use of the provisions by sex exist among all businesses.

Overall, use of these four provisions was low within our study population—less than 7 percent of eligible businesses—based on estimated eligibility in 2020. When analyzed by business size, the largest female-owned S corporations (i.e., in the 75th to 100th quartile) had the highest rate of use for one of the provisions—approximately 16.4 percent used the employer leave credits. All other rates of use were less than 16 percent among the study population by sex and business size.

Estimated Use of COVID-19 Tax Provisions by Small Business Owners' Race and Ethnicity

To estimate race and ethnicity of users of the COVID-19 tax provisions, we imputed race and ethnicity using the address and surname information of taxpayers. We used the Bayesian Improved Surname Geocoding (BISG) imputation method. BISG estimates the probability that individuals having a given surname and residing in a given location identify with each of several racial and ethnic groups (see fig. 10). We estimated these probabilities using data published from the 2010 Census on the race and ethnic identification of individuals living in small geographic areas (i.e., Census Block Groups) and having common surnames. Appendix II describes this method in more detail.

Figure 10: Bayesian Improved Surname Geocoding Imputation Method



Source: GAO analysis of Internal Revenue Service taxpayer data, Social Security Administration data, and U.S. Census Bureau data. | GAO-22-104582

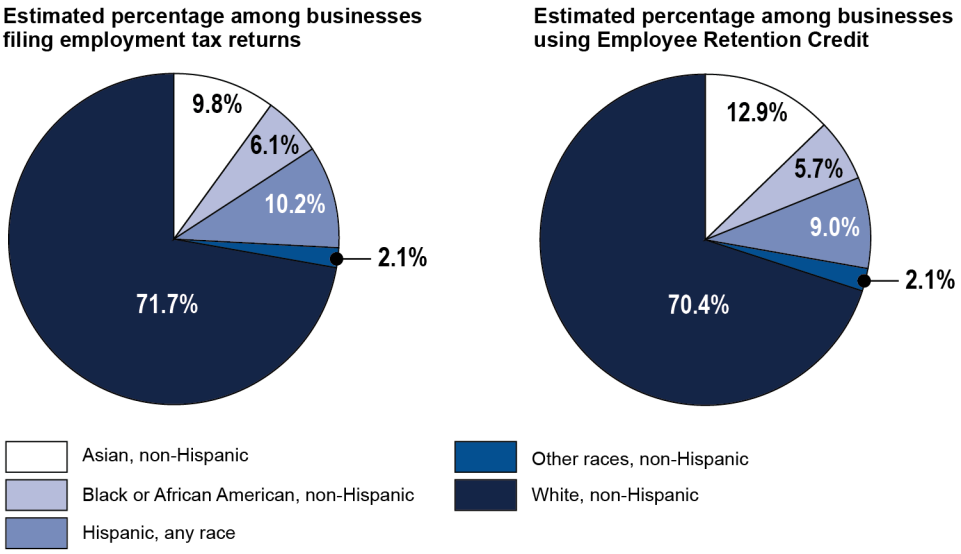
Employee Retention Credit

Among the study population, Asian-owned businesses made up a higher percentage of estimated ERC users, relative to the percentage of Asian-owned businesses filing employment tax returns during tax year 2020 (see fig. 11).[66] Asian-owned businesses filed an estimated 9.8 percent of employment tax returns, but made up an estimated 12.9 percent of ERC users. Conversely, businesses owned by all other race and ethnicity groups were estimated to have lower percentages of ERC users as compared to their percentages among businesses filing employment tax returns. We observed similar patterns across most business size quartiles, though White-owned businesses were slightly overrepresented

[66]For our analysis, White, Black or African American, Asian, and other race groups are of non-Hispanic ethnicity and Hispanic individuals are of any race. "Asian" includes Native Hawaiian and Pacific Islander, non-Hispanic individuals. "Other races" include non-Hispanic American Indian or Alaskan Natives, individuals of two or more races, and individuals in smaller racial groups. The available data sources constrained our measurement of race and ethnicity, as described in appendix II.

among ERC users at the highest quartile of business size (see appendix III, table 15).

Figure 11: Businesses Filing Employment Returns and Use of the Employee Retention Credit by Estimated Race and Ethnicity, among the Study Population, 2020



Source: GAO analysis of Internal Revenue Service taxpayer data, Social Security Administration data, and U.S. Census Bureau data. | GAO-22-104582

Note: Taxpayer data are as reported by taxpayers and subject to taxpayer reporting errors. Use of the Employee Retention Credit is as reported on employment tax forms 941, 943, and 944, and form 7200 for advance payment of the COVID-19 tax provisions. Our analysis included credits claimed on Schedule R and through amended returns associated with form 941. Quarterly returns (forms 941 and 7200) from second through fourth quarter 2020 and annual returns for 2020, including electronically filed returns and paper filings, are included in the figure. The study population included businesses filing employment tax returns, which were also either (1) sole proprietorships filing form 1040, Schedule C or F; or (2) S corporations with a single owner. Data are as of September 2021 to March 2022. “Asian” includes Native Hawaiian and Pacific Islander, non-Hispanic individuals. “Other races” include non-Hispanic American Indian or Alaskan Natives, individuals of two or more races, and individuals in smaller racial groups. Results are not generalizable to the universe of businesses using the provision. Numbers may not add to 100 because of rounding.

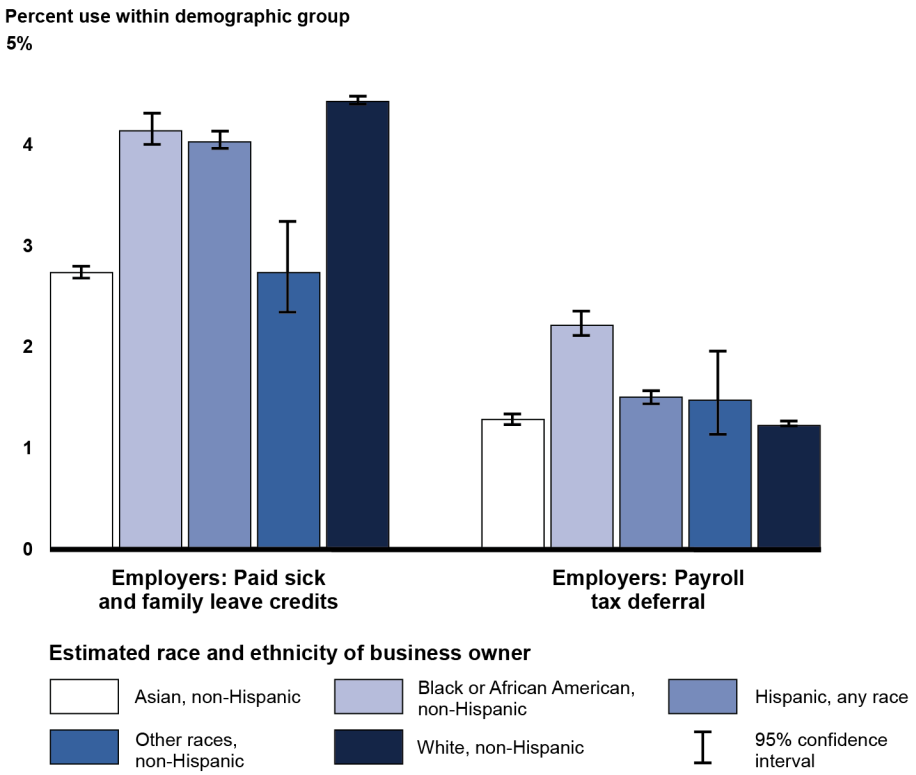
Leave Credits and Payroll Tax Deferrals

For eligible businesses within the study population, we found that the use of selected COVID-19 tax provisions varied by the estimated race and ethnicity of the business owner.[67] For example,

- Black or African American, Hispanic, and White-owned businesses had the highest estimated rates of use for the employer leave credits, with rates of use among these businesses ranging from approximately 4 to 4.4 percent. In contrast, Asian-owned businesses and those owned by individuals of other races both had a lower estimated rate of use at 2.7 percent.
- Black or African American-owned businesses were estimated to be slightly more likely to use the payroll tax deferrals for employers, as compared to businesses owned by all other race and ethnicity groups (see fig. 12).

[67]For the employer paid sick and family leave credits, we used employee counts reported on employment tax forms 941 and 943 to estimate eligibility for the provision. Businesses who reported 500 or more employees in tax year 2020 were excluded from the eligible population. For quarterly returns, we excluded businesses if their average number of employees across the second through fourth quarters of tax year 2020 were 500 or more. Within our study population, an estimated 94.4 to 97 percent of businesses filing employment tax returns were eligible for the leave credits, with eligibility rates varying slightly by demographic group. The number of employees is not collected on form 944, an annual employment return form for the smallest of employers (i.e., businesses whose annual liability for Social Security, Medicare, and withheld federal income taxes is $1,000 or less). Given the small size businesses filing form 944, we assume that all are eligible for the employer leave credits.

Figure 12: Estimated Use of Selected COVID-19 Employer Provisions within Racial and Ethnic Groups, among Eligible Businesses in the Study Population, 2020



Source: GAO analysis of Internal Revenue Service taxpayer data, Social Security Administration data, and U.S. Census Bureau data. | GAO-22-104582

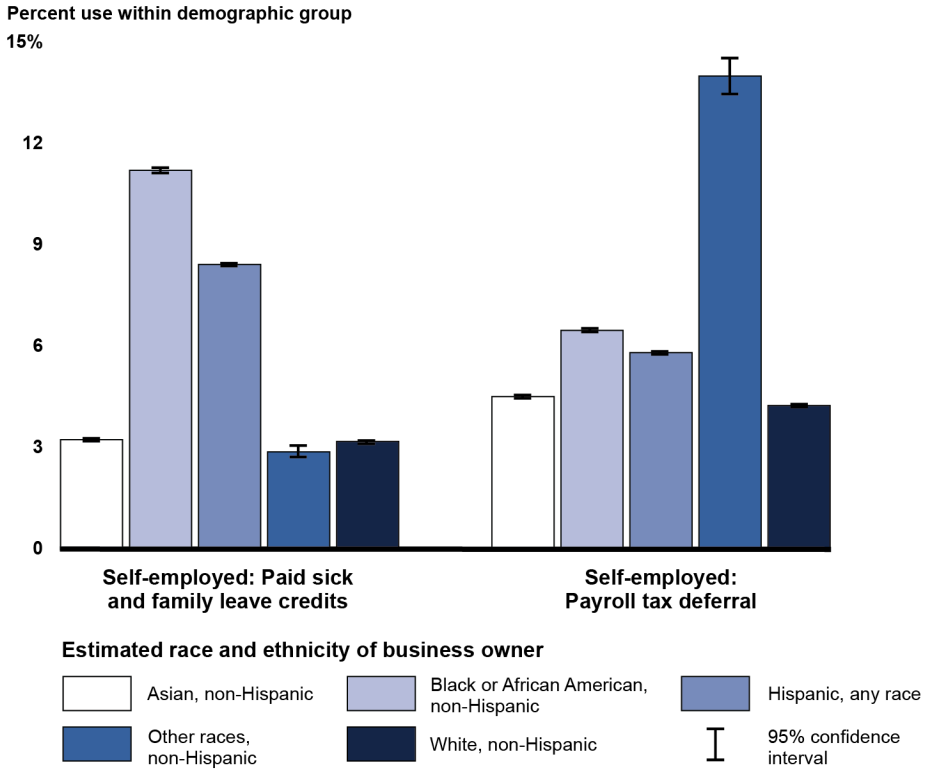
Note: Taxpayer data are as reported by taxpayers and subject to taxpayer reporting errors. Use of the employer provisions is as reported on employment tax forms 941, 943, and 944, and form 7200 for advance payment of the COVID-19 tax provisions. Our analysis included credits claimed on Schedule R and through amended returns associated with form 941. Quarterly returns (forms 941 and 7200) from second through fourth quarter 2020 and annual returns for 2020, including electronically filed returns and paper filings, are included in the figure. The study population included businesses filing employment tax returns and which were also either (1) sole proprietorships filing form 1040, Schedule C or F; or (2) S corporations with a single owner. Data are as of September 2021 to March 2022. Estimates are derived from statistical models, which assumed that the use of tax credits varied around an expected rate for each group, due to random measurement error and other unmeasured variables (see appendix II). We express the uncertainty of these estimates through 95 percent confidence intervals. Results are not generalizable to the universe of businesses using the provisions. "Asian" includes Native Hawaiian and Pacific Islander, non-Hispanic individuals. "Other races" include non-Hispanic American Indian or Alaskan Natives, individuals of two or more races, and individuals in smaller racial groups.

- Black or African American- and Hispanic-owned businesses were more likely to use the leave credits and payroll tax deferral for the self-employed, as compared to Asian- and White-owned businesses. Approximately 11.2 percent of Black or African American-owned businesses and 8.5 percent of Hispanic-owned businesses within our

study population were estimated to use the self-employed leave credits, as compared to 3.3 and 3.2 percent for Asian and White-owned businesses, respectively.

- Businesses owned by other races, which includes Non-Hispanic American Indian or Alaskan Natives, individuals of two or more races, and individuals in smaller racial groups, were estimated to be more likely to use the self-employed payroll tax deferrals, as compared to all other race and ethnic groups (see fig. 13).

Figure 13: Estimated Use of Selected COVID-19 Self-Employed Provisions within Racial and Ethnic Groups, among Eligible Businesses in the Study Population, 2020



Source: GAO analysis of Internal Revenue Service taxpayer data, Social Security Administration data, and U.S. Census Bureau data. | GAO-22-104582

Note: Taxpayer data are as reported by taxpayers and subject to taxpayer reporting errors. Use of the self-employed provisions is as reported on form 1040, Schedule 3. Annual returns for 2020, including electronically filed returns and paper filings, are included in the figure. The study population was sole proprietorships filing form 1040, Schedule C or F, which also filed a form 1040, Schedule SE. Data are as of September 2021 to March 2022. Estimates are derived from statistical models, which assumed that the use of tax credits varied around an expected rate for each group, due to random measurement error and other unmeasured variables (see appendix II). We express the uncertainty of these estimates through 95 percent confidence intervals. Results are not generalizable to the universe of businesses using the provisions. “Asian” includes Native Hawaiian and Pacific Islander,

non-Hispanic individuals. “Other races” include non-Hispanic American Indian or Alaskan Natives, individuals of two or more races, and individuals in smaller racial groups.

These results use statistical models to hold constant business size, as measured by annual receipts reported on tax forms. We obtained similar results when not accounting for business size, though White-owned businesses used the employer leave credits at a slightly higher rate when not accounting for business size (e.g. an estimated 6 percent versus 4.4 percent when controlling for business size). In addition, differences in use by estimated race and ethnicity did not vary substantially by business type (e.g., sole proprietorships and S corporations) for the employer provisions.

While we identified some limitations to imputing race and ethnicity, these estimates indicate that it is feasible to use these or similar methods to analyze race and ethnicity differences at an aggregate level in the use of tax provisions. Further research in this area could yield more complete results. The BISG method has been validated against self-reported racial and ethnic identifications, but only for certain population groups and time periods.[68] The method may not be as effective for specific populations of interest, and may not remain effective over time. Data were unavailable for us to independently validate our estimates. Thus, we do not consider reported estimates to be definitive evidence of differences in use among business owners of varying races and ethnicities. The ability to validate estimates against self-reported race and ethnicity data is a key step in measuring the method’s effectiveness for accurately estimating the demographics of the taxpayer population.

Although we were unable to validate our findings with self-reported race and ethnicity data, we did conduct a sensitivity analysis of our results. Specifically, we ran our analysis on a subset of business owners within our study population. For this subset, the BISG process resulted in a very high probability of belonging to one race group, and very low probability of belonging to all the others. When comparing results from this subset to the original results, we found some noticeable differences in the estimated demographic makeup of businesses filing employment tax returns and those claiming the ERC. These differences were largest for estimations of Black or African American- and White-owned businesses.

[68]Out of the 14 studies we identified in our literature review that used BISG or a variation on the method, 13 studies compared the accuracy of their approach to self-reported data on race and ethnicity. Populations for which these studies examined the method’s accuracy include medical patients, mortgage applicants, business owners, and voters, among others.

However, our overall conclusions based on this high confidence subset were similar to our original results. See tables 10 and 11 in appendix II for full sensitivity analysis results.

The source of our taxpayer data presents additional limitations for this type of analysis. IRS taxpayer data on surnames include only what tax filers report. As a result, the surname data can include various errors by taxpayers, such as suffixes, typographical errors, and inconsistent punctuation. We used the same cleaning methods that Census used to create its surname data to make the surnames as consistent as possible between data sources. However, some inaccuracies likely remain. Additionally, a distinct surname field from a single table in the IRS Compliance Data Warehouse was unavailable for all business owners within the study population. Based on discussions with IRS officials, we used a combination of several fields across different data tables to identify business owner surnames. Using the available cleaned surnames, we matched 84 percent of business owners in the study population to Census surname data. For additional details about our imputation methodology, see appendix II.

Very Small Businesses Struggled to Understand COVID-19 Tax Provisions despite IRS Outreach Efforts

Poor Understanding Contributed to Small Business Owners' Limited Use of COVID-19 Credits and Deferrals

Reasons for Estimated Limited Use

Representatives of almost all of the organizations we interviewed (11 of 12) stated that a lack of understanding could explain why some small

business owners did not use selected credits and deferrals.[69] For our purposes, lack of understanding includes confusion or misunderstanding of the rules, insufficient knowledge of how the provisions work, and difficulty determining eligibility. Representatives from most of the organizations (10) stated that some of the small business owners they work with were confused about their eligibility or lacked sufficient knowledge about the credits and deferrals to claim them.

Representatives from most of the organizations we interviewed (10 of 12) thought small business owners lacked understanding of the COVID-19 tax provisions due in part to difficulty obtaining information from, or working with, IRS or the Small Business Administration (SBA). For example, a representative from one organization suggested that SBA enhance targeted outreach to business owners in low-income communities. Representatives from more than half of these organizations (six of 10) specifically mentioned difficulty obtaining information from or working with IRS. For example, one representative said that many small business owners could not reach anyone at IRS to assist them with questions related to claiming employer credits.[70]

Another representative complimented IRS and SBA on their respective outreach, but acknowledged that it was difficult for agencies to get information to very small businesses.[71] This representative suggested that IRS expand current efforts to publicize available resources, such as seminars and videos, to very small businesses. This interviewee also suggested IRS improve its website by addressing broken links and removing references to deleted pages. Representatives from multiple organizations thought that simplified government communication with taxpayers would be useful, such as one-page documents that provide

[69]We conducted semistructured interviews with 12 organizations in August through November 2021 that represent or work with small business owners. We asked these organizations about owners' experiences using the selected COVID-19 tax provisions, including any challenges small businesses encountered. Additional information on our interviews can be found in appendix I.

[70]As we reported in April 2022, taxpayers had a difficult time reaching IRS during the 2021 filing season due to high call volumes. See GAO, *Tax Filing: 2021 Performance Underscores Need for IRS to Address Persistent Challenges*, GAO-22-104938 (Washington, D.C.: Apr. 11, 2022).

[71]Interviewees who referred to "very small businesses" did not offer definitions in terms of sales or number of employees, but generally described them in the context of businesses lacking resources to have a lawyer or accountant prepare their taxes. Some interviewees referred to "microbusinesses." For purposes of its analysis, SBA describes "microbusinesses" as having one to nine employees.

eligibility and application guidance. For example, a representative from one organization described the guidance for accessing COVID-19 tax provisions as too lengthy and complex for small business owners to understand quickly.

Of the 12 small business organizations we interviewed, representatives from nine also said that insufficient tax preparation resources contributed to a poor understanding of the COVID-19 tax provisions. Our interviewees described these tax preparation resources as tax preparation assistance through a certified public accountant or lawyer. For example, a representative from one organization said that businesses without a certified public accountant often lack the expertise and knowledge to navigate the provisions. Four of our interviewees also suggested that very small businesses are less likely to have resources to access these tax preparation resources than larger businesses.

Complexity of the COVID-19 tax provisions may also be a factor associated with low use. Tax preparation software may alleviate some of the difficulties in determining eligibility and benefit amounts for provisions that require complex computations, according to one study we reviewed.[72] However, the study also states that software cannot reduce the burden on taxpayers in situations where they must provide additional information or keep detailed records to meet the requirements for a tax benefit. Three of the small business organizations we interviewed mentioned that small business owners were confused about requirements for COVID tax provisions even when using tax software.

Our review of the filing instructions for the tax forms associated with the COVID-19 tax provisions indicated that the business owners would need to track specific information. For example, business owners had to track salary information and COVID-related leave hours for each employee to claim the paid leave credits. Business owners also had to track wages used to claim paid leave credits and the ERC since the same wages cannot be used toward both provisions or overlap with Paycheck Protection Program Loan Forgiveness Applications. For those business owners who claimed a payroll tax deferral, they needed to track the amount deferred and the repayment time frame to remain in compliance. In cases where a small business owner has limited recordkeeping

[72]Jacob Goldin, "Tax Benefit Complexity and Take-up: Lessons from the Earned Income Tax Credit," *Tax Law Review,* vol. 72 (2018): 60.

support, it may have been challenging to collect and keep track of this information.

Our review of academic literature identified that knowledge and complexity of tax provisions affected taxpayer behavior.[73] While little research is available on the COVID-19 tax provisions, we reviewed literature on the incomplete take-up of tax benefits generally to understand how knowledge and complexity of a tax benefit might affect taxpayer use. One study found that small and medium businesses did not always claim refunds for prior year tax losses when eligible. Specifically, this study found that the complexity of tax provisions is associated with lower use and professional tax preparation is associated with a higher likelihood of claiming a tax benefit.[74] Another study conducted a field experiment and nongeneralizable survey to assess how outreach materials can affect use of the Earned Income Tax Credit (EITC).[75] The field experiment found that awareness and program complexity can influence taxpayer use of tax provisions, for example by simplifying notices and worksheets. The accompanying survey suggests that these reductions in complexity may have increased credit use by heightening awareness and remedying confusion with respect to eligibility and benefit size.

[73]We reviewed literature about low use of tax provisions in general, but also supplemented our review with research conducted specifically on the use of the COVID-19 tax provisions. See Lucas Goodman, *Take-up of Payroll Tax-Based Subsidies During the COVID-19 Pandemic*, a working paper prepared by staff of the U.S. Department of the Treasury, Office of Tax Analysis, November 2021.

[74]Eric Zwick, "The Costs of Corporate Tax Complexity," *American Economic Journal: Economic Policy*, vol. 13, no. 2 (2021): 497. The study explored the take-up of carryback refunds by analyzing IRS data on U.S. corporate tax return transactions for C corporations between 1998 and 2011.

[75]Saurabh Bhargava and Dayanand Manoli, "Psychological Frictions and Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment," *American Economic Review,* vol. 105, no. 11 (2015): 3518. The study population for the field experiment were individuals from California who filed a 2009 tax return but failed to claim the EITC despite being presumed eligible to receive the credit. The population for the field experiment had two prior opportunities to claim their credit: when they filed their taxes and when receiving an initial reminder. The study population for the field experiment and surveys differs from the broader population of eligible EITC nonclaimants and is nongeneralizable. Further, EITC requirements and eligible populations are different from COVID-19-related provisions. Findings from this study cannot be generalized to other tax benefits or to taxpayers outside the study's population.

Interviewees provided reasons for estimated limited use of the COVID-19 tax provisions among small business owners in addition to poor understanding. Of the 12 organizations we interviewed, representatives from five mentioned lack of awareness—such as not knowing that the program or provision existed—as a reason for limited use of the COVID-19 tax provisions. Further, representatives from four of the 12 organizations said that some small business owners did not use the provisions because they did not find them useful. Organizations that represented Asian, Black or African American, Hispanic, and all types of small business owners also mentioned distrust and fear of government. The representatives from these organizations said small businesses feared making a mistake when claiming a credit and getting into trouble with IRS.

A research paper, government analysis, and our prior work also mention additional possible reasons for limited use of the COVID-19 provisions. A recent study suggested that the Paycheck Protection Program loan may have been preferred over the paid leave credits, since the same wages could not count toward both.[76] This study also mentions the possibility that business owners did not use the leave credits because employees may not have taken leave, or the business may not have been required to provide the leave. The Congressional Research Service reported that the ERC may have been a less beneficial option for employers in 2020 than the Paycheck Protection Program or layoffs, as the credit provided a maximum of $5,000 per employee and could not be used by employers who claimed a Paycheck Protection Program loan.[77] Our May 2022 work also found that retroactive filing requirements and processing delays for

[76]This analysis did not include amended returns and therefore may understate the final take-up of the ERC and paid leave credits. See Goodman, *Take-up of Payroll Tax-Based Subsidies During the COVID-19 Pandemic*, Office of Tax Analysis, 2021. The eligibility change for Paycheck Protection Program borrowers was retroactive to 2020. Thus, employers could file an adjusted employment tax return in 2021 to claim the ERC for qualifying wages paid in 2020. Qualifying wages could not include those used to determine eligibility for the PPP or paid sick or family leave credits. Pub. L. No. 116-260, div. EE, tit. II, § 206, 134 Stat. at 3059–3061.

[77]Congressional Research Service, *The Employee Retention and Employee Retention and Rehiring Tax Credits,* Publication IF11721 (January 2021); and *CARES Act Assistance for Employers and Employees—The Paycheck Protection Program, Employee Retention Tax Credit, and Unemployment Insurance Benefits: Assessment of Alternatives (Part 2)*, Publication IN11329 (April 2020).

the ERC affected the usefulness of the credits, according to tax and payroll professionals.[78]

Variations by Business Size and Demographic Groups

Our interviews indicated there may be differences in understanding the COVID-19 tax provisions based on business size. Representatives from nine of the 12 organizations we interviewed said that owners of very small businesses were less likely than owners of larger small businesses to understand the COVID-19 tax provisions or use professional tax resources. This is consistent with our prior work, which found that eligibility, data collection, and worksheet requirements for the Small Employer Health Tax Credit deterred small employers from claiming it, according to tax preparers, health insurance brokers, and employers.[79]

A recent study measured use of the COVID-19 tax provisions. It found that paid leave credits and payroll tax deferrals were lower for very small businesses, but rose for larger businesses.[80] This study also found that use of the ERC was low overall, although midsize businesses (businesses with approximately 100 employees) were more likely to take the credit than very small or large businesses.

Interviewees from small business organizations who represented particular demographic groups consistently described the same challenges as those faced by small business owners in general. Specifically, representatives from these organizations mentioned poor understanding, difficulty getting information from IRS and SBA, and lack of professional tax preparation assistance. Interviewees also mentioned some challenges that were unique to particular groups, such as language access and technology. For example, a representative from one organization said that many older Asian American and Pacific Islander small business owners do not have access to anyone who can both navigate technology and translate tax information into their native languages. Interviewees from an organization representing Native

[78]GAO-22-104280.

[79]GAO, *Small Employer Health Tax Credit: Factors Contributing to Low Use and Complexity,* GAO-12-549 (Washington, D.C.: May 14, 2012). We interviewed a nongeneralizable sample of tax preparers, health insurance brokers, and employers for this report.

[80]Lucas Goodman, *Take-up of Payroll Tax-Based Subsidies During the COVID-19 Pandemic*, a working paper prepared by staff of the U.S. Department of the Treasury, Office of Tax Analysis, November 2021.

American small business owners mentioned limited broadband internet access in some areas with large indigenous populations.

Businesses owned by certain demographic groups are generally smaller and may have fewer resources to improve understanding of tax provisions. As we discussed previously, small businesses owned by female, Black or African American, and Hispanic individuals tend to have lower average sales than male- or White, non-Hispanic-owned small businesses. The Federal Reserve Banks' 2020 Small Business Credit Survey results reported that nonemployer firms, which are very small businesses, are more likely than employer firms to be woman-, Black or African American-, and Hispanic-owned.[81] Results from the survey also identified that Asian-, Black or African American-, and Hispanic-owned firms were more likely than White-owned firms to report they were in poor financial condition.[82] With limited resources, these small business owners could be less able or willing to pay for professional tax preparation assistance.

IRS Conducted Outreach to Small Businesses, but Lacks Sufficient Evidence to Evaluate Its Outreach

IRS Outreach Efforts

IRS uses its network of partners to share information on tax provisions. IRS officials said that the agency recently expanded its outreach to partners who represent various demographic groups.[83] The IRS *Fiscal Year 2018-2022 Strategic Plan* includes goals to empower and enable taxpayers to meet their tax obligations by improving education and outreach, and to collaborate with partner organizations on services and outreach to taxpayers. Officials told us that IRS does not tailor education and compliance materials based on demographic factors such as race, ethnicity, and sex, but reaches diverse groups by sharing materials through its network of partners. IRS continually expands this partner list,

[81]Federal Reserve Banks, *Small Business Credit Survey: 2021 Report on Nonemployer Firms,* http://www.fedsmallbusiness.org.

[82]Federal Reserve Banks, *Small Business Credit Survey: 2021 Report on Firms Owned by People of Color,* http://www.fedsmallbusiness.org.

[83]According to IRS, it partners with local and national tax practitioner groups, industry associations, advisory groups, congressional offices, and other stakeholders seeking to stay apprised of tax legislation and information. IRS also receives feedback from these partner organizations.

which includes organizations that represent women, multilingual communities, and racial or ethnic groups that have been historically disadvantaged. IRS officials also told us that a focus of outreach efforts have included groups with limited tax resources and access to tax information, such as low-income individuals, veterans, and those experiencing homelessness. According to officials, IRS identifies these new partners through social media platforms and referrals from existing partner organizations.

IRS provides information for small businesses in a variety of ways, including its website, targeted communications, and events. IRS's website provides information on filing taxes specifically for businesses and the self-employed. IRS also distributes an electronic newsletter to small businesses with new information and links to resources on the website. According to IRS officials, IRS participates regularly in tax workshops and national events for small businesses. Additionally, IRS works with SBA district offices and SCORE offices to provide information to very small businesses.[84] Regarding the COVID-19 tax provisions in particular, IRS developed and distributed informational materials to raise awareness. These products generally linked back to information on the IRS website.

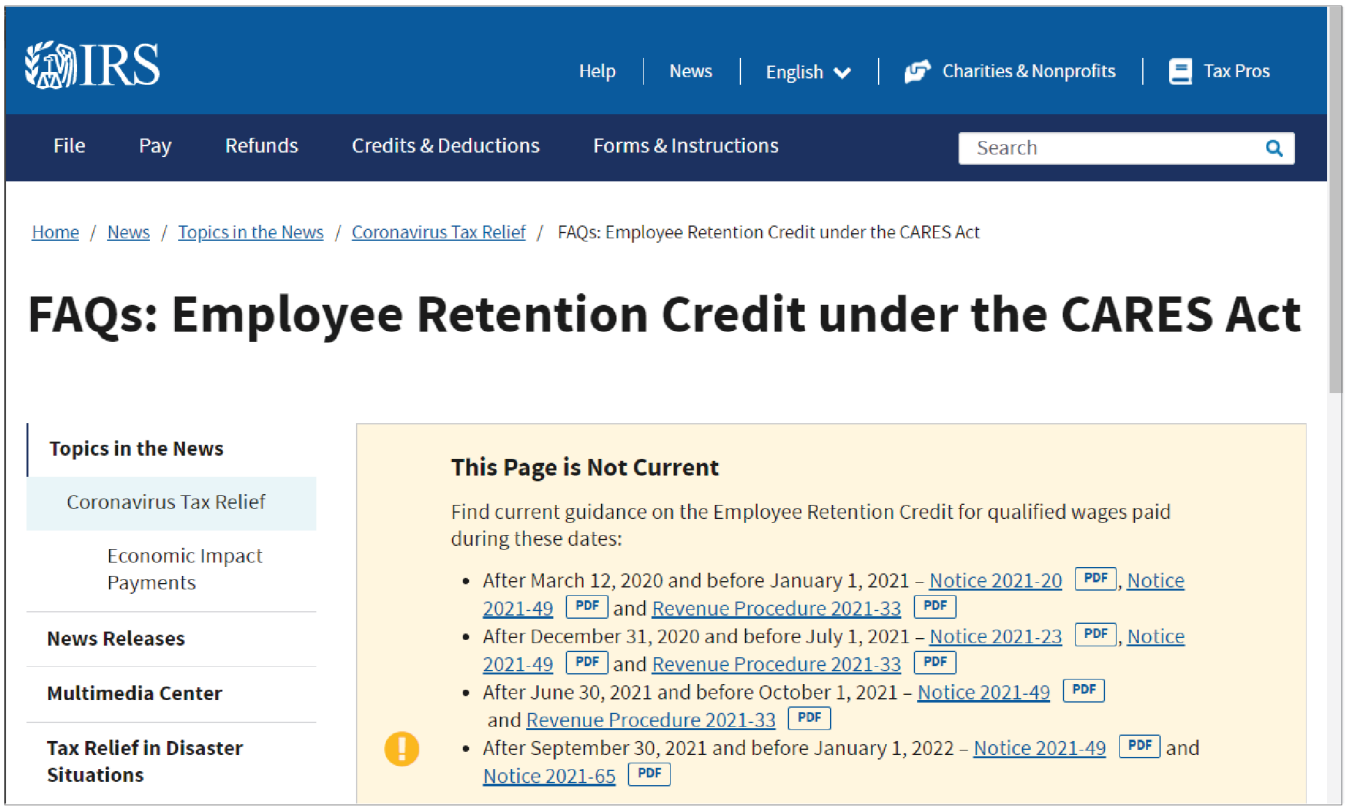
We reviewed information available on the IRS website about the COVID-19 tax provisions, including materials IRS described in its related communications plans. We identified pages where IRS provided tools such as frequently asked questions (FAQ) or visual aids to describe aspects of the COVID-19 provisions. However, we also found instances of pages that were marked as "not current" or "updates underway" and referred visitors to technical guidance documents that reflected changes made with newer laws. For example, the FAQs for the ERC were labeled as "not current" as of April 26, 2022. The web page directed those looking for updated information to multiple guidance documents (see fig. 14). The web page directed those looking for information on qualifying wages after March 12, 2020, and before January 1, 2021, to IRS guidance documents with a combined length of 148 pages.

IRS officials told us that these FAQs were created to provide immediate information to taxpayers while IRS developed formal guidance, and remain active to provide access to the prior information. IRS did not update the FAQs to reflect changes to the law, and instead referred

[84]SCORE is a nonprofit organization that partners with SBA to provide business mentoring and information to small business entrepreneurs.

taxpayers to lengthy and complex guidance documents for updated information. The Taxpayer Advocate's annual report for 2021 cited a lack of proactive transparency and a failure to provide timely, accurate, and clear information to taxpayers as one of the most serious problems facing the agency.[85] The report specifically cited difficulty locating information on IRS's website and unclear IRS guidance as some of the reasons for this lack of transparency.

Figure 14: Screenshot of IRS Website Frequently Asked Questions for the Employee Retention Credit, as of April 26, 2022



Source: https://www.irs.gov/newsroom/faqs-employee-retention-credit-under-the-cares-act. | GAO-22-104582

IRS is working to improve its ability to provide customer support and focused outreach to taxpayers. In March 2022, IRS created a Taxpayer Experience Office to improve taxpayer service. The new office plans to expand customer callback and payment options, offer secure two-way

[85]U.S. Department of the Treasury, Internal Revenue Service, National Taxpayer Advocate, *Annual Report to Congress,* Publication 2104 (December 2021).

messaging, and provide more services for multilingual customers in the short term. These improvements may address the concerns raised by small business organizations regarding difficulty reaching IRS. The 2021 Taxpayer First Act Report to Congress included a taxpayer experience strategy that listed focused outreach to underserved communities as a priority area. This outreach could improve IRS's ability to communicate with various groups, if executed effectively.

IRS Evaluation of Outreach Efforts

Current IRS evaluation of outreach efforts does not provide relevant or complete information to ensure that IRS meets the needs of very small businesses, including those owners from varying demographic groups.

- **Compliance.** IRS officials said they rely on measures of taxpayer compliance to determine the usefulness of its communications. Compliance measures are used to determine if taxpayers pay the appropriate amount of taxes, and to assess the gap between taxes owed and those paid voluntarily and timely. It is unclear how IRS would use compliance measures to determine why those who were eligible did not claim a credit, and how outreach affected their decision.
- **Communication Plans.** IRS's communication plans for the COVID-19 tax provisions included short-term outreach measures such as media attention and web page visits. However, these measures did not evaluate the performance of outreach products for particular groups.
- **Partner Feedback.** IRS officials told us they use informal partner feedback collected through ongoing relationships to refine outreach products. According to IRS officials, IRS staff receive feedback from partners and determine whether to elevate it for response or action. IRS recently established a process to track and analyze the elevated partner feedback for common themes. IRS defines feedback using broad terms such as impressions and perspectives from stakeholders, emerging trends identified by stakeholder liaison staff, or anecdotal evidence. IRS officials said that they meet every quarter to review the feedback and discuss additional actions. However, IRS officials did not describe any efforts to ensure that the feedback is collected systematically or represents the perspectives and needs of different types of small business taxpayers. As a result, IRS does not know if its decisions are based on information received from a limited range of partners. Further, those partners may not sufficiently represent the needs of small businesses that face substantial challenges in understanding tax provisions.

Recent policies have expanded government-wide efforts to address existing inequalities, as well as improve customer experience and service delivery. Executive Order 13985, "Advancing Racial Equity and Support for Underserved Communities Through the Federal Government," directed agencies to "assess whether and to what extent their programs and policies perpetuate systemic inequalities for people of color and other underserved groups."[86] The 2021 President's Management Agenda includes delivering excellent, equitable, and secure federal services as one of its priorities.[87] The Office of Management and Budget (OMB) also recently released guidance directing federal agencies to identify and reduce burdens in accessing public benefits, with a focus on members of underserved and marginalized communities.[88]

The period of eligibility has passed for the COVID-19 tax provisions. Modifications to associated outreach materials would only assist those claiming credits retroactively through amended returns.[89] However, evaluating ongoing and future outreach efforts using relevant and complete information could help IRS improve its communications to various groups more generally, and plan for future emergencies.

In its January 2021 report to Congress on the Taxpayer First Act, IRS said it planned to use data to identify necessary language translations for its notices and correspondence, and to improve communication with taxpayer groups who face unique challenges accessing tax information. While the report describes strategies to develop focused outreach to underserved communities, it does not include specific objectives or measures to evaluate outreach to small businesses. IRS officials we spoke to did not indicate any plans to change how they evaluate outreach to small businesses.

Treasury has integrated equity goals into its fiscal year 2022-2026 strategic plan, including a strategy to increase outreach, education, and

[86]Exec. Order No. 13985, 86 Fed. Reg. 7009 (Jan. 25, 2021).

[87]President's Management Council and the Office of Management and Budget, *The Biden-Harris Management Agenda Vision* (Nov. 18, 2021).

[88]Office of Management and Budget, *Improving Access to Public Benefits Programs Through the Paperwork Reduction Act,* OMB Memorandum M-22-10 (Washington, D.C.: Apr. 13, 2022).

[89]Business owners may be able to claim COVID-19 tax credits through amended returns, which generally must be filed within 3 years after the date of the original return.

compliance tools for underserved communities.[90] Implementing the Treasury equity goals could help address differences in access to information among groups. IRS should use measurable objectives and collect relevant and complete data to evaluate its efforts. Evaluating outreach efforts for very small businesses and those owners with various demographic backgrounds would allow IRS to determine if its performance meets Treasury's and government-wide strategic goals. This evaluation can inform and improve IRS outreach in ways that are useful to those in need of assistance. If IRS were able to improve evaluation of its outreach efforts for particular groups, it could better understand whether communication needs for these groups are being met. This information could inform outreach efforts in future emergencies and help improve small business owners' understanding of tax provision eligibility.

Conclusions

The COVID-19 tax provisions helped some employers maintain payroll and address the health-related leave needs of employees and self-employed business owners. However, we found that very small business owners may have struggled to use the credits and deferrals due, in part, to a poor understanding of the provisions. Representatives from small business organizations we interviewed reported challenges in getting clear information from IRS and SBA, and difficulty accessing professional tax preparation assistance as reasons for this poor understanding. While other reasons for low use of the provisions were also mentioned by representatives from the small business organizations we interviewed, a poor understanding of the provisions was mentioned most frequently and by organizations representing varying demographic groups of business owners. Our review of the relevant tax forms also identified complexities associated with claiming the provisions, such as detailed information businesses would need to track.

We found that IRS could improve its ability to convey complicated and rapidly changing information about tax provisions if it used relevant and complete information to evaluate outreach efforts against measurable objectives. Improved evaluation of outreach could also help IRS refine outreach efforts during future emergencies, and alleviate barriers for small business owners, including those with varying racial and ethnic backgrounds. Treasury has efforts underway to estimate use of tax

[90]Department of the Treasury, *Treasury Strategic Plan 2022-2026* (Washington, D.C.: 2022).

provisions by demographic groups. This could also help refine future education and outreach.

Recommendation for Executive Action

The Commissioner of IRS should evaluate IRS's outreach efforts to very small businesses and owners with diverse backgrounds, using relevant and complete information, to inform future outreach. (Recommendation 1)

Agency Comments

We provided a draft of this report to the Secretary of the Treasury, Commissioner of the Internal Revenue Service, Secretary of Commerce, and Administrator of the Small Business Administration for review and comment. IRS agreed in principle with our recommendation and its comments are reproduced in appendix V. We also received technical comments from Treasury, IRS, and SBA, which we incorporated as appropriate. Commerce said it had no comments on the draft report.

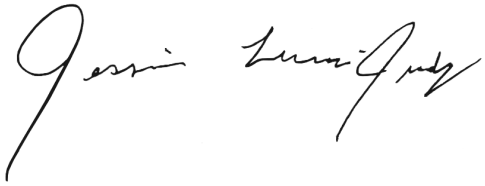
In its letter, IRS agreed that it should evaluate its outreach efforts and emphasized the complexities associated with reaching diverse groups of small business owners using demographic data, since such data are not tracked or collected by IRS. IRS also stated that it plans to collaborate with Treasury on demographic data collection, which it will use to refine future education and outreach materials.

We acknowledge the complexities IRS faces in evaluating outreach to different groups in the absence of demographic data on tax provision use. We agree that working with Treasury to estimate use of tax provisions by demographic groups could help refine future outreach efforts. While demographic data is one information source, IRS could also consider using other information to enhance outreach evaluation, such as equity measures built into the partner feedback process. Such efforts could help IRS refine its outreach during future emergencies and alleviate barriers for small business owners, including those with varying racial and ethnic backgrounds.

We are sending copies of this report to the appropriate congressional committees, the Secretary of the Treasury, Secretary of Commerce, Commissioner of Internal Revenue Service, and the Administrator of the Small Business Administration. In addition, the report is available at no charge on the GAO website at https://www.gao.gov.

If you or your staff have any questions about this report, please contact me at (202) 512-6806 or lucasjudyj@gao.gov. Contact points for our

Offices of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made key contributions to this report are listed in appendix VI.



Jessica Lucas-Judy
Director, Tax Issues
Strategic Issues

List of Committees

The Honorable Patrick Leahy
Chairman
The Honorable Richard Shelby
Vice Chairman
Committee on Appropriations
United States Senate

The Honorable Ron Wyden
Chairman
The Honorable Mike Crapo
Ranking Member
Committee on Finance
United States Senate

The Honorable Patty Murray
Chair
The Honorable Richard Burr
Ranking Member
Committee on Health, Education, Labor, and Pensions
United States Senate

The Honorable Gary C. Peters
Chairman
The Honorable Rob Portman
Ranking Member
Committee on Homeland Security and Governmental Affairs
United States Senate

The Honorable Rosa L. DeLauro
Chair
The Honorable Kay Granger
Ranking Member
Committee on Appropriations
House of Representatives

The Honorable Frank Pallone, Jr.
Chairman
The Honorable Cathy McMorris Rodgers
Republican Leader
Committee on Energy and Commerce
House of Representatives

The Honorable Bennie G. Thompson
Chairman
The Honorable John Katko
Ranking Member
Committee on Homeland Security
House of Representatives

The Honorable Carolyn B. Maloney
Chairwoman
The Honorable James Comer
Ranking Member
Committee on Oversight and Reform
House of Representatives

The Honorable Richard E. Neal
Chairman
The Honorable Kevin Brady
Republican Leader
Committee on Ways and Means
House of Representatives

Appendix I: Objectives, Scope, and Methodology

This report (1) describes the distribution of small business owners by race, ethnicity, and sex; (2) describes information and methods available to examine differences in the use of selected tax provisions by demographics of small business owners; (3) estimates the use of COVID-19 tax provisions by demographics of small business owners for a selected study population; and (4) evaluates potential barriers in accessing COVID-19 tax provisions among small businesses, including whether these barriers varied by demographic group.

Distribution of Small Business Owners by Race, Ethnicity, and Sex

To describe the distribution of small business owners by race, ethnicity, and sex, we analyzed data from multiple U.S. Census Bureau datasets. We analyzed 2019 data—the most recent year available—on business characteristics from the Annual Business Survey (ABS) to describe demographics of employer businesses.[1] We also analyzed 2018 data—the most recent year available—from the Nonemployer Statistics by Demographics (NES-D) data series to describe the distribution of demographics for nonemployer small businesses.[2]

We were unable to restrict our analysis of employer businesses to small businesses (commonly defined as having fewer than 500 employees) due to data suppression issues. Census suppresses some employer data estimates (1) to avoid disclosing data for individual businesses, and (2) if it does not meet publication standards because of high sampling variability, poor response quality, or other concerns about the estimate quality. However, we found that most employer businesses (an estimated

[1]Census' ABS is an electronic survey that sampled approximately 300,000 employer businesses in 2020. The most recent data are the result of the 2020 ABS survey, but use 2019 as the reference year. Business ownership is defined as having 51 percent or more of the stock or equity in the business. Businesses are categorized by (1) race (White, Black or African American, American Indian and Alaska Native, Asian, or Native Hawaiian and Other Pacific Islander); (2) ethnicity (Hispanic, equally Hispanic and non-Hispanic, or non-Hispanic); and (3) sex (male, female, or equally male and female). Not all businesses are classifiable by race, ethnicity, and sex. Businesses can be counted in more than one race group if the sole owner, majority owner, or majority combination of owners was reported to be of more than one race. We calculated percentages for employer businesses using the total number of classifiable businesses as the denominator.

[2]Census developed the NES-D data series to produce similar estimates as ABS on owner demographics for nonemployer businesses. The NES-D is not a survey; rather, it leverages existing individual-level administrative records to assign demographic characteristics to the universe of nonemployer businesses. Not all businesses are classifiable by race, ethnicity, and sex. We calculated percentages for nonemployer businesses using the total number of classifiable businesses as the denominator. Business ownership definitions and race, ethnicity, and sex categories for NES-D are similar to ABS.

99.7 percent) have less than 500 employees. Therefore, we use the term "small business" for all employer estimates in this report. We also considered all nonemployer businesses to be small businesses for our purposes.

We also used the employer and nonemployer data to examine business demographics by sector.[3] In prior work, we used information from the 2020 Bureau of Labor Statistics' Business Response Survey to identify the six "hardest-hit sectors," or those most likely to experience adverse effects to business operations as a result of the pandemic.[4] These six sectors are (1) accommodation and food services; (2) arts, entertainment, and recreation; (3) educational services; (4) health care; (5) manufacturing; and (6) retail trade.

To determine if business ownership was proportional to the demographic characteristics of the population, we analyzed 2019 data from Census' American Community Survey (ACS) 5-year estimates.[5]

For the employer and community survey data analysis, we used reported standard errors to calculate 95 percent confidence intervals around estimates. We determined that differences between estimates were statistically significant if the 95 percent confidence interval surrounding one estimate did not overlap with the 95 percent confidence interval

[3]Sector classification in the employer and nonemployer data is based on the North American Industry Classification System, which is the standard used by federal statistical agencies in classifying business establishments according to industry.

[4]Adverse effects to business operations included a shortage of supplies or inputs, decreased demand for products or services, difficulty moving or shipping goods, and government-mandated closure of a business location. See GAO, *Paycheck Protection Program: Program Changes Increased Lending to the Smallest Businesses and Underserved Locations,* GAO-21-601 (Washington D.C.: Sept. 21, 2021).

[5]Census defines race as a person's self-identification with one or more social groups. An individual can report as White, Black or African American, Asian, American Indian and Alaska Native, Native Hawaiian and Other Pacific Islander, or some other race. ACS respondents may report multiple races. For purposes of comparison with employer and nonemployer data, we used totals that represented the maximum number of individuals who reported as that race group, either alone, or in combination with another race(s). We also combined some groups and reported on four race groups: Asian, Black or African American, White, and Other (which includes American Indian and Alaska Native as well as Native Hawaiian and Other Pacific Islander). We did not include those who reported "some other race" in our calculations. According to Census, ethnicity or origin can be viewed as the heritage, nationality group, lineage, or country of birth of the person or the person's parents or ancestors before their arrival in the United States. Individuals who identify their origin as Hispanic, Latino, or Spanish may be of any race. Census defines sex as biological sex with two possible responses, female or male.

surrounding the other estimate. For the nonemployer data, Census rounds the number of businesses and introduces "noise" or distortions into the data on sales amounts as the primary method of disclosure avoidance. Based on Census' information on rounding and noise flags, we calculated an interval for each estimate that contains the population value. We consider differences between estimates to be significant if the interval surrounding one estimate did not overlap with the interval surrounding the other estimate.

To assess the reliability of the employer, nonemployer, and community survey datasets, we reviewed technical documentation on methodology and reviewed responses from knowledgeable agency officials. We determined the data used in our analyses were sufficiently reliable for the purposes of this report.

Information and Methods Available to Examine Differences in Use of Selected Tax Provisions by Demographics of Small Business Owners

We reviewed data that the federal government collects on the race, ethnicity, and sex of small business owners. We reviewed information from all 13 federal statistical agencies and two additional agencies that collect data on small businesses. We identified datasets for this objective through systematic searches of selected federal agency websites using search terms such as "business" and "demographics," reviews of relevant research, and interviews with agency officials. We included publicly reported data that selected federal agencies collected in our review. Data were determined to be relevant for our purposes if it included information on business ownership as well as race, ethnicity, or sex information. Selected federal agencies included the principal federal statistical agencies, as well as agencies with missions that include collecting data on small businesses.[6] Additionally, we reviewed technical documentation that described collection methodology and limitations for each dataset we included in our review.

We also reviewed relevant studies to identify analytical methods for assigning race, ethnicity, and sex when that information is missing from a dataset. To identify studies, we searched various databases including

[6]The Office of Management and Budget identified 13 federal statistical agencies, including the Bureaus of Economic Analysis, Justice Statistics, Labor Statistics, and Transportation Statistics; Census Bureau; Department of Agriculture, Economic Research Service; Energy Information Administration; Internal Revenue Service, Statistics of Income; National Agricultural Statistics Service; the National Centers for Education Statistics, and Health Statistics; National Science Foundation, National Center for Science and Engineering Statistics; and Social Security Administration, Office of Research, Evaluation, and Statistics. Selected federal agencies that collect data on small businesses include the Small Business Administration and the Federal Reserve System.

Scopus, ProQuest, EBSCO, and Harvard Think Tank Search using search terms related to imputation, estimation, administrative data, race, ethnicity, and sex. We identified 132 studies that appeared in scholarly publications, working papers, government reports, and publications by associations, nonprofits, and think tanks over the last 15 years. These studies were relevant to our research objective on methods to identify missing race, ethnicity, and sex data. We performed searches in March and September 2021; through team research, we identified an additional relevant working paper that was released after our literature searches in October 2021.

We limited our review to studies that applied methods that could be relevant to analyzing taxpayer data. We used the following criteria to review summary information on each study to determine its relevance:

- Studies published in peer-reviewed journals over the last 10 years;
- Studies containing original research-based findings; and
- Studies that applied a method for obtaining missing race, ethnicity, or sex information on an administrative dataset.[7]

We identified 39 studies that met our inclusion criteria. We then conducted a full-text review of each study to collect detailed information and assess strengths and limitations of each method. We reviewed information for each study such as population, research objectives, methodology, data sources, data analysis techniques, and variables. Using an inductive approach with independent coders, we developed the classification scheme and used it to categorize missing data methods. We iteratively refined and tested the classification scheme, and documented it in a codebook. We conducted our review using a standardized data collection instrument. We eliminated one study that we determined was neither methodologically sound nor rigorous enough for our purposes. We also included one working paper published after our database searches that was relevant and sufficiently reliable for our purposes. Appendix IV lists the 39 studies included in our review.

To describe agency efforts to analyze tax provisions by taxpayer demographics we interviewed officials at Census, the Internal Revenue Service (IRS), and the Department of the Treasury.

[7]"Last 10 years" included articles published in 2011 through 2021.

Estimated Use of COVID-19 Tax Provisions by Demographics of Small Business Owners

To examine use of selected COVID-19 tax provisions by race, ethnicity, and sex of small business owners, we analyzed IRS taxpayer data, Social Security Administration (SSA) data, and publically available Census data. Selected tax provisions included three COVID-19 tax provisions for employers: the Employee Retention Credit (ERC), paid sick and family leave credits (leave credits), and the deferrals of tax payments for the employer's share of payroll tax. Self-employed business owners were also eligible for the leave credits and payroll tax deferral. We analyzed use of the leave credits and payroll tax deferrals separately for employers and self-employed business owners. Complete data were unavailable for tax year 2021 at the time of our analysis. Therefore, our analysis covers use of these provisions on 2020 annual returns and on quarterly returns for the second through fourth quarters of 2020. We consider a business to have used a provision if it reports any nonzero dollar amount in the related tax return field.

Selected Study Population

Assigning demographics to businesses with multiple owners is difficult and requires complex decision rules on determining shares of ownership and assigning demographics to the business unit accordingly. For ease of analysis, we limited businesses within scope to those with a single owner. Our selected study population of small businesses included (1) sole proprietorships filing form 1040, *U.S. Individual Income Tax Return*, Schedule C, *Profit or Loss from Business (Sole Proprietorship*), or Schedule F, *Profit or Loss from Farming*; and (2) S corporations where a single individual was identified across ownership forms.[8] We excluded partnerships and C-corporations because these business forms are more likely to have complex ownership structures and ownership that is difficult to determine based on tax returns. Our study population did not include all users of the selected COVID-19 tax provisions. In addition, because we could not select users with known probabilities, our analysis did not produce results that generalize to the universe of businesses using the provisions.

[8]Data on ownership of S corporations were obtained from Form 1120-S, Schedule K-1, *Shareholder's Share of Income, Deductions, Credits, etc.* Corporations use Schedule K-1 to report individuals' shares of the corporation's income, deductions, credits, and other items. We excluded S corporations where a Form 1120-S did not have a corresponding Schedule K-1. Approximately 94 percent of 1120-S returns in 2020 have a corresponding Schedule K-1. We could not determine what portion of the missing Schedule K-1s are associated with single-owner S corporations.

To the extent possible, we excluded sole proprietorships operated as a qualified joint venture by a married couple.[9] To remove qualified joint ventures, we identified married couples who filed joint income returns but separate Schedule Cs, and who each listed a Schedule C business with the same industry and business description.[10] We required that the Schedule C business descriptions matched exactly. We also excluded businesses filing Schedule C or Schedule F with a nonunique employer identification number (EIN).[11]

We further narrowed our selected study population based on the particulars of the provisions we analyzed. For analysis of the employer provisions, our study population consisted of approximately 2.8 million businesses filing an employment tax return.[12] This represents approximately 39 percent of all employment tax returns in the second through fourth quarters of 2020. EINs were used to match employment tax returns to income and business tax returns. For analysis of the self-employed provisions, our study population consisted of approximately 15.4 million sole proprietorships with a Form 1040, *Individual Income Tax Return,* Schedule SE*, Self-Employment Tax*, in 2020.[13] These businesses

[9]According to an IRS website, a qualified joint venture is one that conducts a trade or business where (1) the only members of the joint venture are a married couple who file a joint return, (2) both spouses materially participate in the trade or business, and (3) both spouses elect not to be treated as a partnership. http://www.irs.gov/businesses/small-businesses-self-employed/election-for-married-couples-unincorporated-businesses.

[10]Due to data availability issues, we were unable to take this same approach to identifying qualified joint ventures among businesses filing Form 1040, Schedule F.

[11]Businesses do not have to be employers to obtain an EIN. We identified a number of businesses filing Form 1040, Schedule C (sole proprietorship) that use the same EINs. Through discussions with IRS, we determined that these sole proprietors are unlikely to be employers; instead, they are using an EIN associated with a larger company, most likely in error. To ensure that we are connecting the EIN with the correct business owner, we exclude filers using nonunique EINs on Schedules C or F.

[12]This count of employment tax returns included unique EINs across Form 941, *Employer's Quarterly Federal Tax Return*; Form 943, *Employer's Annual Federal Tax Return for Agricultural Employees;* and Form 944, *Employer's Annual Federal Tax Return.* It also included businesses that used Form 941, Schedule R, to file employment tax returns. Data are as of September 2021 to March 2022.

[13]We did not include S corporations in our analysis of the self-employed provisions because S corporations pay employment taxes through employment tax returns rather than self-employment tax returns.

are owned by 14.2 million business owners, representing 78 percent of all Schedule SE filings in 2020.[14]

Provision Eligibility Considerations

Not all businesses within our study population were eligible for each selected tax provision. To represent use of the provisions accurately by demographic group, we estimated an eligible population of businesses for almost all provisions, using available information within the taxpayer data and based on IRS guidance for employers and self-employed business owners.

- **Employer sick and family paid leave credits.** We estimated businesses filing employment tax returns to be eligible for the leave credits if their number of employees in tax year 2020 were less than 500.[15] We obtained the employee count data from Form 941, *Employer's Quarterly Federal Tax Return* and Form 943, *Employer's Annual Federal Tax Return for Agricultural Employees*.[16] Within our study population, about 96 percent of businesses filing employment tax returns were eligible for the leave credits. The number of employees is not collected on Form 944, an annual employment return form for the smallest of employers (i.e., businesses whose

[14]According to IRS, generally someone is self-employed if they (1) carry on a trade or business as a sole proprietor or an independent contractor; (2) are a member of a partnership that carries on a trade or business; or (3) are otherwise in business for themselves (including a part-time business). Because of our focus on analyzing demographics of eligible provision users, we did not apply any tests based on income or deductions to narrow the population to those engaged in business activity, as it is traditionally understood. Our study population is based on tax forms that report business income and the unit of analysis is businesses rather than individuals. Thus, we refer to the self-employed as "business owners" in discussion of the results.

[15]This is an estimation of eligibility due to differences between the concept of number of employees as captured on the tax forms and the definition for purposes of the leave credits. For example, Form 941 asks for the employee count as of a specific pay period in the quarter, while eligibility for leave credits is determined by employee counts on the dates that the employees took leave. For quarterly returns, we estimated businesses to be eligible if their average number of employees across the second through fourth quarters of tax year 2020 were less than 500.

[16]In March 2021, we found that some employers using the leave credits may be ineligible for the credits based on the number of employees reported on their Forms 941. IRS officials said that they had identified taxpayer errors in some employers' entries and, based on our recommendation, issued a "tax tip" in May 2021 for employment tax return filers, reminding them to ensure that employee counts on their return are accurate. See GAO-21-387. Based on our review of the data, we determined that employee counts were sufficiently reliable for our specific purpose of calculating eligibility for the leave credits. Our analysis comparing use of the leave credits by all users within the study population, to eligible users within the study population, showed similar results by demographic groups.

annual liability for Social Security, Medicare, and withheld federal income taxes is $1,000 or less). Given the small size businesses filing form 944, we assume that all are eligible for the employer leave credits.

- **Employer payroll tax deferral.** We assumed that all businesses filing employment tax returns within our study population were eligible to defer the employer share of payroll taxes.
- **Self-employed sick and family leave credits and payroll tax deferral.** We assumed that all businesses filing Form 1040, Schedule SE, were eligible for both the leave credits and payroll tax deferral.

We were unable to determine eligibility for the ERC due to eligibility rule complexity and lack of data on quarterly receipts and government-ordered suspension of operations. For the ERC, we compared the demographic makeup of businesses filing employment tax returns within our study population to that of businesses using the ERC.

To control for the effects of business size in our analysis, we examined annual business receipts—obtained from tax returns for 2020 when estimating how usage rates varied across demographic groups.[17]

Provision Use Analysis

To describe use of the selected COVID-19 tax provisions, we analyzed IRS data from Forms 941, *Employer's Quarterly Federal Tax Return*; Form 943, *Employer's Annual Federal Tax Return for Agricultural Employees*; Form 944, *Employer's Annual Federal Tax Return*; Form 7200, *Advance Payment of Employer Credits Due to COVID-19*; and Form 1040, *U.S. Individual Income Tax Return*, Schedule 3, *Additional*

[17]Tax return fields used for receipts analysis varies by form. For Form 1040, Schedule C and 1120-S, we analyzed gross receipts data from lines 1 and 1a, respectively. For Form 1040, Schedule F, there are multiple lines for reporting various receipts and income sources and not all taxpayers report values in all these fields. Therefore, we analyzed the field that was most populated—the gross income from the cash method, which is reported on line 9. For some businesses filing a Form 1120-S, 2020 receipts data were unavailable.

Credits and Payments.[18] We also included adjusted tax returns for businesses filing 941-X. Third-party payers, such as a payroll company filing returns on behalf of many clients, use Schedule R to allocate their clients' dollar amounts for each line on Form 941. We included data on the individual employers listed on Schedule R.

Our data may not include all 2020 tax returns because there were time lags in when IRS uploads the data. IRS stores each form's data in a different database. Thus, the dates covered are different for each form (see table 1).

Table 1: Tax Return Forms Analyzed and Dates of Data Updates

Return	Date data updated
Form 941	January 24, 2022
Form 941, Schedule R	September 20, 2021 to the week of March 11, 2022
Form 941-X (adjusted returns)	November 15, 2021
Form 943	January 24, 2022
Form 944	January 24, 2022
Form 7200	Week of March 11, 2022
Form 1040, Schedule 3	December 31, 2021
Form 1040, Schedule C	January 24, 2022
Form 1040, Schedule F	January 24, 2022
Form 1040, Schedule SE	January 24, 2022
Form 1120-S	January 24, 2022
Form 1120-S, Schedule K-1	January 24, 2022

Source: GAO analysis of Internal Revenue Service taxpayer data. | GAO-22-104582

The tax provision use we are reporting are subject to taxpayer reporting error. We report what was filed without adjustments. For data from

[18]We count a business filing an employment tax return as having used a provision if the business, as identified through its unique EIN, reported use on any one Form 941 or 7200 in the second through fourth quarters of 2020 or on Form 943 or 944 in 2020. If a business claimed the same provision across multiple forms or for multiple quarters in 2020, we count that as one use. Self-employed individuals file for their leave credits and report payroll tax deferrals on their income tax return. We matched Schedule 3 data on provision use to study population businesses by filer's Social Security number. We also limited our analysis to returns with a Schedule SE to exclude provision use by filers of Schedule H, *Household Employment Taxes*. Some self-employed business owners (approximately 7.5 percent of the 14.2 million owners studied) had more than one business that fell within our study population during tax year 2020. In these cases, the provision use is counted more than once.

adjusted returns, such as Form 941-X, we analyzed account transaction codes and credit reference numbers to identify employers that used the provisions. Unlike taxpayer-provided data on employment tax returns, these codes reflect actions on the employer's account.

Data Matching to Identify Sex of Business Owners

For our study population of businesses, we matched owner Social Security numbers, as provided on the relevant tax forms, to SSA data to identify the individual's recorded sex. We accessed SSA data on sex of individuals through IRS's Compliance Data Warehouse. These data are updated monthly and were accessed on March 1, 2022. According to SSA officials, these data originate with the *Application for a Social Security Card*, completed for all Social Security number holders. SSA officials said that current SSA systems only allow for female or male sex entries, but there may be some "unknown" entries in the sex field if the sex was unknown on older paper applications that were later converted to electronic records.

We matched 99.96 percent of business owners within our study population as either female or male. We use the terms female and male because the dataset we analyze uses these terms to describe sex variables. Furthermore, our analysis does not consider gender, as it signifies social and cultural factors not present in our data. See table 2 for the sex of business owners among our study population.

Table 2: Percentage of Businesses in Study Population by Sex of Owner, 2020

	Businesses filing employment tax returns	Businesses filing self-employment tax returns
Female	27.45%	34.02%
Male	72.53%	65.94%
Unknown or missing sex	0.03%	0.04%

Source: GAO analysis of IRS taxpayer and Social Security Administration data. | GAO-22-104582.

Note: Employment tax returns include forms 941, 943, and 944, and form 7200 for advance payment of the COVID-19 tax provisions. Our analysis included form 941, Schedule R returns. For employment tax returns the study population consisted of businesses that were either (1) sole proprietorships filing form 1040, Schedule C or F; or (2) S corporations with a single owner. A self-employment tax return is one that included form 1040, Schedule SE. For the self-employed provisions, the study population included sole proprietorships filing form 1040, Schedule C or F. Quarterly returns (forms 941 and 7200) from second through fourth quarter 2020 and annual returns for 2020, including electronically filed returns and paper filings, are included in the table. Data are as of September 2021 to March 2022.

Imputation to Estimate Race and Ethnicity of Business Owners

We used a prediction method, the Bayesian Improved Surname Geocoding method, to estimate race and ethnicity, using IRS and SSA data on business owners' surname and home address as well as publicly available Census data. For details on how we used this method, see appendix II.

Data Reliability

We assessed the reliability of Census, IRS, and SSA data by reviewing relevant documentation, interviewing knowledgeable agency officials, and performing electronic testing. We determined that the data used in our analysis were sufficiently reliable for the purpose of estimating use of the provisions by demographic groups of business owners.

Potential Barriers in Accessing COVID-19 Tax Provisions among Small Businesses

To evaluate potential barriers for accessing COVID-19 tax provisions among small businesses, and whether these barriers varied by demographic group, we conducted semistructured interviews with 12 organizations that represent or work with small business owners. We conducted these interviews in August through November 2021. We asked representatives of these organizations about owners' experiences using the selected COVID-19 tax provisions, including any challenges small businesses encountered. We identified organizations through interviewee recommendations (a snowball sample), a literature search, and other research on relevant organizations. We then determined those organizations most appropriate to interview by evaluating publicly available information to determine that each organization:

- was relevant to the group it represents or advocates for;
- interacted with small business owners that would allow it to speak to the broad experiences of the businesses it represented; and
- was active during the COVID-19 pandemic.

To ensure we heard a range of perspectives from different organization types (e.g., nonprofits and membership organizations) and relevant to varying demographic groups (e.g., organizations representing Asian, Black or African American, Hispanic, or Native American small business owners and women small business owners), the team used a random stratified selection process. Specifically, organizations were sorted into categories based on organization type and whom they represented. At least one organization was then randomly selected within each group, when possible. If an organization declined or did not respond to our interview request, the team selected another organization using the same random stratified sample until we interviewed at least one organization

representing each category. The interviews provided anecdotal information that is not generalizable to all organizations.

We also searched for literature that discussed take-up of federal tax provisions to better understand the factors described in our interviews and find additional factors. We identified eight studies published in the last 10 years initially through a search of the Scopus, ProQuest, ProQuest Dialog, and EBSCO databases in October 2021. We then supplemented our literature search results by adding four potentially relevant studies that we independently identified in January 2022.[19] The criteria we used to determine relevance for studies we included in our analysis included take-up of tax provisions, primary factors described in our interviews with organizations, and relevance to small business owners. We also conducted a high-level fatal flaw review to determine whether the study's limitations called the study's findings into question. Of the seven studies that met our criteria, we removed one during a fatal flaw analysis. In total, our literature review included six studies.

To understand how IRS addressed barriers to accessing tax benefits when administering the COVID-19 provisions, we collected information on its outreach and communication activities. To do this, we interviewed IRS officials and reviewed agency documentation on its outreach to small businesses on the COVID-19 provisions. We also collected information on how IRS evaluates its education and outreach activities and reviewed relevant IRS policies and procedures. We then compared IRS's evaluation of outreach efforts to standards in Executive Order 13985, "Advancing Racial Equity and Support for Underserved Communities Through the Federal Government," and the *Treasury 2022-2026 Strategic Plan*.[20]

We conducted this performance audit from October 2020 through August 2022 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe

[19]We chose supplemental studies by identifying relevant studies and reviewing works cited in the studies. We also included an article (Goodman, 2021) that was released after our database search due to the relevance of the topic (take-up rates for COVID-19 tax provisions).

[20]Exec. Order No. 13985, 86 Fed. Reg. 7009 (Jan. 25, 2021); Department of the Treasury, *Treasury Strategic Plan 2022-2026* (Washington, D.C.: 2022).

that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Appendix II: Methods for Estimating Race and Ethnicity

We estimated the race and ethnicity of tax filers within our scope using the Bayesian Improved Surname Geocoding (BISG) method.[1] BISG predicts identification with selected racial and ethnic groups using a filer's surname and residential location. We view our estimation as a demonstration of feasibility, not necessarily as a definitive estimate of racial disparities, given limitations in the method we describe in the body of this report.

Literature Review and Prior Use of BISG

We reviewed the literature on racial and ethnic imputation methods—and specifically those using surnames—published since 2006 to assess methods that could be applied with Internal Revenue Service (IRS) tax data. We conducted systematic database searches to identify literature. Our electronic search covered several databases of peer-reviewed scholarly publications, conference papers, working papers, government reports, and publications by associations, nonprofits, and think tanks.[2]

We concluded that the BISG method has been widely used and feasible to use with taxpayer data. Other methods of imputation exist for specific datasets that have self-reported racial or ethnic data and auxiliary variables unique to the dataset, such as insurance enrollment records or custom surveys, which tax data do not contain. In contrast, BISG can estimate race and ethnicity using only surname and residential address, which tax data do contain.

Our literature review found 18 studies that validated BISG estimates against self-reported racial and ethnicity identification collected through administrative or survey sources in the healthcare, financial, and other sectors. This research has found that BISG accurately predicts race and ethnicity for individuals identifying as Hispanic and non-Hispanic White, Black or African American, or Asian/Pacific Islander. For example:

[1]Marc N. Elliott, et al., "Using the Census Bureau's Surname List to Improve Estimates of Race/Ethnicity and Associated Disparities," *Health Services and Outcomes Research Methodology*, vol. 9, no. 69 (2009): 69-83.

[2]The purpose of this review was to inform our imputation methodology. We included studies in our review that developed or evaluated indirect estimation or imputation methods for predicting race or ethnicity. Although similar substantively, the purpose and scope of this review differed from the literature review discussed in the body of our report used to answer our second research objective. Specifically, there were some differences in search parameters to identify literature and criteria to select literature. For more information on the purpose and scope of the literature review discussed in the body of our report, see appendix I.

- The original developers of BISG reported correlations between predicted probabilities and self-reported values for the four groups above (combining all Asians) ranging from .70 to .82, using 2006 national healthcare insurance plan enrollment data, but reported .01 to .11 correlations for American Indian/Alaskan Native and multiracial groups. A similar study by the same developers found correlations for these groups ranging from .61 to .79.[3]
- A later independent study found that BISG predicted self-reported racial and ethnic identifications from administrative and survey datasets with 76 to 96 percent accuracy, depending on the validation dataset. However, BISG predicted identification with other groups less accurately, such as for American Indian and Alaskan Native, and multiracial identifiers; predictive accuracy rates for these groups were 6 to 42 percent.[4]
- Two validation studies of BISG found that the method can perform less well among women than among men, possibly due to higher rates of surname changes among women.[5]

We determined that BISG was sufficient to demonstrate a potentially valid method to estimate race and ethnicity using IRS taxpayer data—the goal of our analysis. Our literature review identified other imputation methods, such as variations of BISG that used first name and age, that could be applied to other datasets, such as voter registration and Medicare and Medicaid beneficiary data. Some of these other methods required variables that were unavailable in taxpayer data or used data from sources on selected subpopulations that would have required substantial efforts to assess reliability. As a result, we selected BISG as the method that (1) predicted race and ethnicity with acceptable accuracy, and (2) used available taxpayer data and reliable surname and demographic lookup data from the U.S. Census Bureau. Future efforts to estimate

[3]Marc N. Elliott, et al., "A New Method for Estimating Race/Ethnicity and Associated Disparities Where Administrative Records Lack Self-Reported Race/Ethnicity," *Health Services Research*, vol. 43, no. 5, Part I (October 2008): 1730-1732.

[4]Derose, et al., "Race and Ethnicity Data Quality and Imputation Using U.S. Census Data in an Integrated Health System: The Kaiser Permanente Southern California Experience," *Medical Care Research and Review*, vol. 70, no. 3 (2012): 338-339.

[5]Adjaye-Gbewonyo, et al., "Using the Bayesian Improved Surname Geocoding Method (BISG) to Create a Working Classification of Race and Ethnicity in a Diverse Managed Care Population: A Validation Study," *Health Services Research*, vol. 49, no. 1, Part I (February 2014): 278-279. Caroline K. Smith and David K. Bonauto, "Improving Occupational Health Disparity Research: Testing a Method to Estimate Race and Ethnicity in a Working Population." *American Journal of Industrial Medicine*, vol. 61, no. 8 (August 2018): 643–644.

taxpayer race and ethnicity, outside of our demonstration, could explore alternative methods or data sources.

Data

Two public use data files from the 2010 Census of Population and Housing allowed us to apply the BISG method to tax data. When we designed our analysis, Census had not published the equivalent files for the most recent 2020 census.

2010 Census Surname File

The 2010 Census Surname File contains aggregate data on the probability that a respondent having a given surname reported race and ethnicity in one of six categories:

- Hispanic or Latino, Any Race;
- Non-Hispanic or Latino, White Alone;
- Non-Hispanic or Latino, Black or African American Alone;
- Non-Hispanic or Latino, Asian, Native Hawaiian, or Pacific Islander Alone;
- Non-Hispanic or Latino, American Indian or Alaskan Native Alone; and
- Non-Hispanic or Latino, Two or More Races.

BISG developers reported poor validation performance for the latter two groups. So, we combined them into a residual "Other/Unknown" group by subtracting the sum of the probabilities across all the other reported groups from one.[6]

A distinct surname field from a single table in the IRS Compliance Data Warehouse was unavailable for all business owners within the study population. Depending on the source, some data fields contain full names for individuals while others contain concatenated names of married couples filing jointly. Based on discussions with IRS officials, we used a combination of several fields across different data tables to identify a surname for each business owner. These data tables included surname information collected on tax forms and by the Social Security Administration (SSA).[7] IRS surname information was updated between January 24 and February 21, 2022. We accessed this information on

[6]Elliott, et al., "Using the Census Bureau's Surname List," 77.

[7]One data table includes surnames of individuals who have made changes to their information with SSA. These changes could include, for example, if someone is issued a new Social Security number.

March 1, 2022. Surname information from SSA was updated between December 20, 2021, and February 21, 2022. We accessed this information on March 1, 2022.

We reviewed written responses from knowledgeable IRS officials to determine which data tables to use. This process assigns a name to 94.7 percent of all business owners in our study population. See tables 3 and 4 for number and percent of business owners within the study population with surnames assigned and matched to Census data. We edited these names using the same methods employed by Census in creating the Census Surname file. These edits removed common suffixes and other unreliable data. For business owners associated with a full name, as opposed to a surname only, we used the last word in the name string after the edits were applied as the surname.

Table 3: Number and Percent of Business Owners with Surname Assigned by Tax Form, 2020

	Number with surname assigned	Number of total records	Percent with surname assigned
Form 1040 Schedule C filers	27,983,082	29,388,928	95.22%
Form 1040 Schedule F filers	1,417,465	1,716,080	82.6%
Form 1120-S single-owner filers	3,060,549	3,161,240	96.81%
Total business owners within the study population	**32,461,096**	**34,266,248**	**94.73%**

Source: GAO analysis of Internal Revenue Service taxpayer and Social Security Administration data. | GAO-22-104582

Table 4: Number and Percent of Business Owners with Surname Matched to Census Surname File by Tax Form, 2020

	Number with surname matched to Census surname file	Number of total records	Percent with surname matched to Census surname file
Form 1040 Schedule C filers	24,840,209	29,388,928	84.52%
Form 1040 Schedule F filers	1,335,656	1,716,080	77.83%
Form 1120-S single-owner filers	2,625,012	3,161,240	83.04%
Total business owners within the study population	**28,800,877**	**34,266,248**	**84.05%**

Source: GAO analysis of Internal Revenue Service taxpayer, Social Security Administration, and U.S. Census Bureau data. | GAO-22-104582

Census included only the surnames that at least 100 respondents provided. These surnames covered 90 percent of the individuals with surnames recorded in the 2010 Census. Census reports that 95.5 percent of respondents provided a surname, implying that the surnames in the file covered about 86 percent of respondents overall. To further preserve

privacy, Census suppressed counts for racial and ethnic groups making up a small proportion of respondents with a given surname. Following the developers of BISG, we imputed the suppressed counts for missing cells as *n/k*, where *n* is the sum of the counts in the suppressed groups and *k* is the number of suppressed groups.[8] The probability imputed for each cell was then *(n/k) / N*, where *N* is the total count for the surname.

2010 Census Summary File 1

The 2010 Census Summary File 1 (SF1) is a public use file of statistics on population and housing for various geographic areas. BISG has been developed and validated using racial data for block groups, which are groups of street blocks. Accordingly, we filtered the SF1 to include statistics on block groups in all 50 states and the District of Columbia.

We selected statistics on the number of adults in each block group who identified with the five racial groups in the Census 2010 Surname File, combining "Asian alone" and "Hawaiian and Other Pacific Islander Alone" and calculating an equivalent "other" group. Both the SF1 and surname files used data from the 2010 Census. So, they used the same methods to define racial groups and measure identification. We used aggregate counts for each block group to calculate the proportion of each block group that identified with each racial group.

Estimation of Tax Filer Race and Ethnicity

We applied the BISG method as described in the literature, but changed selected steps to reflect unique features of the tax data in our scope.[9]

The IRS tax return data in our scope contained the surnames, residential addresses, and use of tax credits for N tax filers, $i = \{1, \ldots, N\}$. Let $T = 1$ if the filer claimed a credit and 0 if they did not. The filers reported one of $s = \{1, \ldots, 162{,}253\}$ surnames and reported addresses in one of $g = \{1, \ldots, 216{,}684\}$ 2010 Census block groups. Each filer identified with one of five unknown racial groups defined above (including "other"), $r = \{0, \ldots, 4\}$.

We identified block groups from addresses by applying the geocoding algorithms in the SAS geocode procedure. This procedure maps taxpayer-supplied street addresses from tax forms to two-digit state, three-digit county, six-digit tract, and one-digit block group Federal Information Processing Standards (FIPS) codes using lookup data based

[8]Elliott, et al., "Using the Census Bureau's Surname List," 73.

[9]Elliott, et al., "Using the Census Bureau's Surname List." Kosuke Imai and Kabir Khanna, "Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Records," *Political Analysis*, vol. 24, no. 2 (Spring 2016): 263-272.

on Census geographic shape files. We then concatenated these to create the 12-digit FIPS codes that match the geographic IDs from the Census SF1 file.

We performed minimal cleaning on the taxpayer-supplied addresses prior to running the geocode process. Specifically, apartment numbers and suite numbers were removed. While street addresses containing p.o. boxes were neither removed nor edited, 39 percent of addresses that were not assigned a matching FIPS code included the string “PO BOX.” Taxpayer address information was available in two tables, which were updated on January 24, 2022, and February 21, 2022, respectively. We accessed these tables on March 1, 2022.

IRS address data are more complete than surname data. More than 99.9 percent of business owners in our study population had an assigned address. The geocode process produced block groups that matched to the Census SF1 file for 88.7 percent of all business owners in the study population. See tables 5 and 6 for number and percent of business owners within the study population with valid residential addresses and geocoded block group matched to Census data. See table 7 for the number and percent of business owners within the study population with both a geocoded block group and surname matched to Census data.

Table 5: Number and Percent of Business Owners with a Valid Residential Address by Tax Form, 2020

	Number with address assigned	Number of total records	Percent with address assigned
Form 1040 Schedule C filers	29,380,863	29,388,928	99.97%
Form 1040 Schedule F filers	1,716,072	1,716,080	100%
Form 1120-S single-owner filers	3,160,683	3,161,240	99.98%
Total business owners within the study	**34,257,618**	**34,266,248**	**99.97%**

Source: GAO analysis of Internal Revenue Service taxpayer data. | GAO-22-104582

Table 6: Number and Percent of Business Owners with Geocoded Block Group Matched to Census SF1 File by Tax Form, 2020

	Number with block group assigned and matched to Census SF1 file	Number of total records	Percent with block group assigned and matched to Census SF1 file
Form 1040 Schedule C filers	26,186,137	29,388,928	89.1%
Form 1040 Schedule F filers	1,403,264	1,716,080	81.77%
Form 1120-S single-owner filers	2,793,366	3,161,240	88.36%
Total business owners within the study	**30,382,767**	**34,266,248**	**88.67%**

Source: GAO analysis of Internal Revenue Service taxpayer and U.S. Census Bureau data. | GAO-22-104582

Table 7: Number and Percent of Business Owners with Geocoded Block Group Matched to Census SF1 File and Surname Matched to Census Surname File by Tax Form, 2020

	Number with name matched to Census surname file AND block group assigned and matched to Census SF1 file	Number of total records	Percent with name matched to Census surname file AND block group assigned and matched to Census SF1 file
Form 1040 Schedule C filers	22,140,843	29,388,928	75.34%
Form 1040 Schedule F filers	1,090,995	1,716,080	63.57%
Form 1120-S single-owner filers	2,314,655	3,161,240	73.22%
Total business owners within the study	**25,546,493**	**34,266,248**	**74.55%**

Source: GAO analysis of Internal Revenue Service taxpayer, Social Security Administration, and U.S. Census Bureau data. | GAO-22-104582

We estimated Pr(T = 1 | R = r), the population probability of claiming a tax credit, given identification with one of several race and ethnic groups. We used the following method[10]:

1. We estimated the probability that the filer identified with each racial group, given their surname, using the 2010 Census surname file: Pr(R_i = r | S_i = s). For filers with surnames that did not match the Census list, we attempted to identify whether the name was hyphenated. We then assigned a racial probability if both components of the name matched the same racial group (e.g., Hernandez-Perez but not Smith-Hernandez). For any other surname that did not match,

[10]Elliott, et al., "Using the Census Bureau's Surname List." Kosuke Imai and Kabir Khanna, "Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Records," *Political Analysis*, vol. 24, no. 2 (Spring 2016): 263-272.

we assigned the probabilities for the “ALL OTHER NAMES” entry in the surname file.

2. We estimated the probability that the filer identified with each racial group, given their 2010 Census block group, using the 2010 Census SF1: Pr(*R_i* = *r* | *G_i* = *g*). If the block group was missing, we set the final racial imputation for each group as the probability linked to the filer’s surname, per step 1.
3. We estimated the probability that the filer lived in each 2010 Census block group using the 2010 Census SF1: Pr(*G_i* = *g*).
4. We assumed that geolocation and surname were statistically independent, given race or ethnicity. This means that filers with the same race or ethnicity, having different surnames, are equally likely to live in the same block group.
5. We applied Bayes’ Rule to estimate the probability of living in a block group, *g*, given a racial identification, *r*, by substituting the variables in the prepped files for each quantity:

$$\Pr(G_i = g|R_i = r) = \frac{\Pr(R_i = r|G_i = g) \cdot \Pr(G_i = g)}{\sum_g \Pr(R_i = r|G_i = g) \cdot \Pr(G_i = g)}$$

   where the sum is taken over all block groups.

6. Finally, we estimated race or ethnicity by applying Bayes’ Rule again to estimate the probability that filer *i* identified with racial group *r,* given their surname s and block group *g*:

$$P_{ir} = \Pr(R_i = r|S_i = s, G_i = g) = \frac{\Pr(G_i = g|R_i = r) \cdot \Pr(R_i = r|S_i = s)}{\sum_r \Pr(G_i = g|R_i = r) \cdot \Pr(R_i = r|S_i = s)}$$

   where the sum was taken over all racial groups for each filer. The imputation yielded a set of five probabilities that the tax filer identified with each racial group.

Descriptive statistics on the estimated probabilities appear in tables 8 and 9, separately by businesses filing employment or self-employment tax returns and by the type of missing data. The mean probability is the estimated percentage of the sample identifying with a group. The quantiles are the percentages of the sample having estimates below the listed values. For example, among businesses filing employment tax returns, an estimated 10.4 percent of the sample with surnames and Census block groups identified as Asian (non-Hispanic). In addition, 90 percent of the sample had estimated probabilities below 52.3 percent. As

expected, missing data on either surname or Census block group affected the estimates because the other observed covariate exclusively predicted race and ethnicity. These differences affect about 27 percent of the estimation sample.

Table 8: Imputed Race and Ethnic Probabilities for Businesses Filing Employment Tax Returns by Type of Missing Data

	Matched surname and Census block group	Missing surname	Missing Census block group
N	2,032,004	487,023	327,536
Asian, non-Hispanic			
Mean	10.4%	9.0%	7.6%
10th quantile	0	0.5	0.4
25th quantile	0.1	1.7	0.5
50th quantile	0.3	5.6	0.7
75th quantile	1.2	10.1	5.7
90th quantile	52.3	22.2	8
Black or African American, non-Hispanic			
Mean	5.8	4.7	8.5
10th quantile	0	0.2	0.2
25th quantile	0	0.5	0.6
50th quantile	0.3	1.6	8.5
75th quantile	2.8	6.3	16.1
90th quantile	15.6	8.7	28.4
Hispanic, any race			
Mean	10.3	9.7	10.7
10th quantile	0.1	1	1.5
25th quantile	0.2	2	2.2
50th quantile	0.6	5.1	2.5
75th quantile	2.2	13.7	11.8
90th quantile	41.3	20.6	13.7
White, non-Hispanic			
Mean	71.5	74.1	68.5
10th quantile	1.7	44.3	6.7
25th quantile	50.8	66.7	65.9
50th quantile	93	79.1	72.9

	Matched surname and Census block group	Missing surname	Missing Census block group
75th quantile	97.9	90.9	91.3
90th quantile	99.1	95.9	95.1

Source: GAO analysis of Internal Revenue Service taxpayer, Social Security Administration, and U.S. Census Bureau data. | GAO-22-104582

Note: Entries are probabilities for businesses filing employment tax returns and which were also either (1) sole proprietorships filing form 1040, Schedule C or F; or (2) S corporations with a single owner in 2020. Qualified joint ventures and businesses with nonunique EINs have been removed. Results are not generalizable to the universe of businesses. N = 2,791,139. Estimates for a residual group, “Other (non-Hispanic),” are not reported.

Table 9: Imputed Race and Ethnic Probabilities for Businesses Filing Self-Employment Tax Returns by Type of Missing Data

	Matched surname and Census block group	Missing surname	Missing Census block group
N	11,289,800	2,681,806	1,694,656
Asian, non-Hispanic			
Mean	7.1%	8.9%	6.3%
10th quantile	0	0.4	0.4
25th quantile	0	1.4	0.5
50th quantile	0.2	5.1	0.7
75th quantile	0.9	9.7	5.2
90th quantile	6.8	22.7	8
Black or African American, non-Hispanic			
Mean	11.2	7.2	11
10th quantile	0	0.2	0.3
25th quantile	0	0.7	0.5
50th quantile	0.5	2.4	8.5
75th quantile	6.2	8.5	16.1
90th quantile	44.5	16.4	30.7
Hispanic, any race			
Mean	22.1	14.2	19.6
10th quantile	0.1	1.1	1.7
25th quantile	0.3	2.5	2.3
50th quantile	1	7.4	2.7
75th quantile	13	14.3	13.7
90th quantile	97.8	37.8	92

	Matched surname and Census block group	Missing surname	Missing Census block group
White, non-Hispanic			
Mean	57.5	66.9	60.5
10th quantile	0.3	21.7	4.9
25th quantile	4.4	54	48.7
50th quantile	80.1	71.4	66.7
75th quantile	96.5	89.1	86.2
90th quantile	98.8	95.5	94.6

Source: GAO analysis of Internal Revenue Service taxpayer, Social Security Administration, and U.S. Census Bureau data. | GAO-22-104582

Note: Entries are probabilities for sole proprietorships filing form 1040, Schedule C or F, and which also filed a form 1040, Schedule SE in 2020. Qualified joint ventures and businesses with nonunique EINs have been removed. Results are not generalizable to the universe of businesses. N = 15,365,483. Estimates for a residual group, "Other (non-Hispanic)," are not reported.

Analyzing Tax Credit Usage by Estimated Race and Ethnicity

We estimated the probability that a filer would claim a tax credit, given the filer's estimated probabilities of identifying with the racial and ethnic groups, *p_i*, using the following logistic regression model:

$$E(T_i|p_i) = \frac{\exp(p_i\beta)}{1 + \exp(p_i\beta)}$$

where *piβ* is a linear combination of the four estimated racial identification probabilities, excluding the probability of identifying as White (the baseline comparison group), and their coefficients. The model included an intercept for the White baseline comparison group (including 1 in the first position of the *p_i* vector). We estimated probabilities for each group using the estimated coefficients and setting the group's covariate to 1 and the values for all other groups to 0.[11] Finally, we calculated the 95 percent confidence interval of this predicted probability for reporting purposes.

We assessed the sensitivity of our estimates to two methodological choices. First, we replicated selected estimates of employment tax returns and Employee Retention Credit use by race, using a subsample where the imputed probabilities exceeded 90 percent, without controlling for size, as shown in table 10. Second, we assessed the sensitivity of

[11]The literature recommends using the predicted probabilities as covariates, rather than classifying each filer into a single group, to avoid losing statistical power when using the estimates. See Daniel F. McCaffrey and Marc N. Elliott, "Power of Tests for a Dichotomous Independent Variable Measured with Error," *Health Services Research*, vol. 43, no. 3 (June 2008): 1085-1101.

estimates across groups to unmeasured differences in size by including covariates measuring the quartiles of firm revenue, as shown in table 11. We estimated probabilities for each group as above, separately for all estimates and those that exceeded 90 percent, except that we averaged the calculations over the sample distribution of the size covariate.

Table 10: Sensitivity Results for Businesses Filing Employment Returns and Use of the Employee Retention Credit by Estimated Race and Ethnicity, among the Study Population, 2020

	High-confidence estimations		All estimations	
	Estimated percent among businesses filing employment tax returns	Estimated percent among businesses using Employee Retention Credit	Estimated percent among businesses filing employment tax returns	Estimated percent among businesses using Employee Retention Credit
Asian, non-Hispanic	9.79%	13.82%	9.84%	12.92%
Black or African American, non-Hispanic	1.88	1.72	6.14	5.68
Hispanic, any race	9.06	7.3	10.2	8.97
Other races, non-Hispanic	1.18	1.15	2.12	2.06
White, non-Hispanic	78.09	76.02	71.71	70.36

Source: GAO analysis of Internal Revenue Service taxpayer, Social Security Administration, and U.S. Census Bureau data. | GAO-22-104582

Note: Taxpayer data are as reported by taxpayers and subject to taxpayer reporting errors. Use of the Employee Retention Credit is as reported on employment tax forms 941, 943, and 944, and form 7200 for advance payment of the COVID-19 tax provisions. Our analysis included credits claimed on Schedule R and through amended returns associated with form 941. Quarterly returns (forms 941 and 7200) from second through fourth quarter 2020 and annual returns for 2020, including electronically filed returns and paper filings, are included in the table. The study population included businesses filing employment tax returns, which were also either (1) sole proprietorships filing form 1040, Schedule C or F; or (2) S corporations with a single owner. Data are as of September 2021 to March 2022. Results are not generalizable to the universe of businesses using the provision. “Asian” includes Native Hawaiian and Pacific Islander, non-Hispanic individuals. “Other races” include non-Hispanic American Indian or Alaskan Natives, individuals of two or more races, and individuals in smaller racial groups. Numbers may not add to 100 because of rounding.

Table 11: Sensitivity Results for Estimated Use of COVID-19 Tax Provisions within Racial and Ethnic Groups, among the Study Population, 2020

		High-confidence estimations		All estimations	
Tax provision	Estimated race and ethnicity of the business owner	Estimated percent use within the demographic group	95 percent confidence interval (lower bound, upper bound)	Estimated percent use within the demographic group	95 percent confidence interval (lower bound, upper bound)
Employers: Paid sick and leave credits	Asian, non-Hispanic	2.81%	(2.73, 2.89)	2.74%	(2.67, 2.8)
	Black or African American, non-Hispanic	5.11	(4.79, 5.46)	4.14	(3.99, 4.3)
	Hispanic, any race	4.25	(4.14, 4.36)	4.03	(3.95, 4.12)
	Other races, non-Hispanic	3.52	(2.35, 5.23)	2.74	(2.33, 3.22)
	White, non-Hispanic	4.47	(4.42, 4.51)	4.43	(4.39, 4.46)
Employers: Payroll tax deferral	Asian, non-Hispanic	1.25%	(1.19, 1.3)	1.29%	(1.24, 1.34)
	Black or African American, non-Hispanic	2.47	(2.25, 2.7)	2.22	(2.1, 2.35)
	Hispanic, any race	1.52	(1.45, 1.58)	1.51	(1.46, 1.57)
	Other races, non-Hispanic	1.47	(0.72, 2.98)	1.48	(1.13, 1.96)
	White, non-Hispanic	1.22	(1.2, 1.24)	1.23	(1.21, 1.24)
Self-employed: Paid sick and leave credits	Asian, non-Hispanic	3.63%	(3.59, 3.68)	3.26%	(3.22, 3.3)
	Black or African American, non-Hispanic	10.99	(10.89, 11.09)	11.24	(11.17, 11.32)
	Hispanic, any race	8.55	(8.51, 8.59)	8.45	(8.42, 8.49)
	Other races, non-Hispanic	3.79	(3.32, 4.33)	2.9	(2.73, 3.09)
	White, non-Hispanic	3.07	(3.06, 3.09)	3.21	(3.2, 3.23)
Self-employed: Payroll tax deferral	Asian, non-Hispanic	4.37%	(4.32, 4.42)	4.54%	(4.49, 4.58)
	Black or African American, non-Hispanic	5.93	(5.85, 6)	6.5	(6.45, 6.56)
	Hispanic, any race	5.71	(5.67, 5.74)	5.84	(5.81, 5.86)
	Other races, non-Hispanic	11.48	(10.62, 12.41)	14.03	(13.51, 14.56)
	White, non-Hispanic	4.05	(4.04, 4.07)	4.27	(4.25, 4.28)

Source: GAO analysis of Internal Revenue Service taxpayer, Social Security Administration, and U.S. Census Bureau data. | GAO-22-104582

Note: Taxpayer data are as reported by taxpayers and subject to taxpayer reporting errors. Use of the employer provisions is as reported on employment tax forms 941, 943, and 944, and form 7200 for advance payment of the COVID-19 tax provisions. Our analysis included credits claimed on Schedule R and through amended returns associated with form 941. Use of the self-employed provisions is as reported on form 1040, Schedule 3. Quarterly returns (forms 941 and 7200) from second through fourth quarter 2020 and annual returns for 2020, including electronically filed returns and paper filings, are included in the table. For the employer provisions, the study population included businesses filing employment tax returns, which were also either (1) sole proprietorships filing form 1040, Schedule C or F; or (2) S corporations with a single owner. For the self-employed provisions, the study population included sole proprietorships filing form 1040, Schedule C or F, and which also filed a form 1040, Schedule SE. Data are as of September 2021 to March 2022. Estimates are derived from statistical models, which assumed that the use of tax credits varied around an expected rate for each group, due to random measurement error and other unmeasured variables. We express the uncertainty of these estimates through 95 percent confidence intervals. Results are not generalizable to the universe of businesses using the provisions. “Asian” includes Native Hawaiian and Pacific Islander, non-Hispanic individuals. “Other races” include non-Hispanic American Indian or Alaskan Natives, individuals of two or more races, and individuals in smaller racial groups.

Appendix III: Data Analysis Supplemental Tables

Table 12: Ownership Rates in Hardest-Hit Sectors by Race, Ethnicity, and Sex of Business Owners

	Total all sectors	Accommodation and food services	Arts, entertainment, and recreation	Educational services	Health care and social assistance	Manufacturing	Retail trade
Nonemployers							
Asian-owned	8.04	11.78*	4.31*	6.73*	8.57*	5.13*	6.93*
Black or African American-owned	12.07	19.11*	9.68*	10.04*	22.45*	7.19*	8.24*
White-owned	79.23	68.22*	85.65*	82.92*	68.34*	86.8*	84.25*
Owned by individuals of other races	0.5	0.64*	0.47*	0.42*	0.59*	0.57*	0.47*
Hispanic-owned	15.1	19*	8.48*	8.92*	15.73*	13.78*	11.82*
Non-Hispanic-owned	84.7	80.89*	91.39*	91.02*	84.22*	85.92*	87.98*
Equally Hispanic and non-Hispanic-owned	0.19	0.21*	0.11*	0.05*	0.05*	0.29*	0.16*
Female-owned	42.08	51.78*	40.25*	62.47*	75.84*	32.84*	56.64*
Male-owned	55.4	45.78*	58.88*	37.03*	23.72*	64.22*	41.72*
Equally male and female-owned	2.51	2.44	0.83*	0.44*	0.43*	2.79*	1.62*
Employers							
Asian-owned	10.45	25.86*	3.4*	11.17	13.53*	4.68*	16.88*
Black or African American-owned	2.42	1.56*	3.56	3.77*	6.75*	0.69*	1.41*
White-owned	86.63	71.59*	92.62*	84.36	78.96*	94.07*	81.28*
Owned by individuals of other races	0.6	0.54	0.36	S	0.63	0.51	S
Hispanic-owned	6.23	8.41*	3.58*	4.64*	5.64	4.84*	4.87*
Non-Hispanic-owned	92.85	90.43*	95.07	94.05	93.62*	94.41*	94.49*
Equally Hispanic and non-Hispanic-owned	0.92	1.16*	1.35	S	0.75	0.75	0.65*
Female-owned	21.72	22.23	21.17	44.15*	34.43*	16.47*	21.56
Male-owned	63.06	57.14*	61.64	38.85*	55.3*	68.28*	59.05*
Equally male and female-owned	15.22	20.64*	17.19*	17	10.26*	15.26	19.39*

Source: GAO analysis of U.S. Census Bureau data. | GAO-22-104582

Note: Nonemployer data are as of 2018; employer data are as of 2019. (*) indicates that differences between sector-level ownership rates and ownership rates across all sectors are significant at the 95 percent confidence interval for employer businesses and based on calculated intervals for nonemployer businesses. All employer estimates have a margin of error of ±4 percent or less at the 95 percent confidence level, with some exceptions. Margins of error were greater for estimates of White- (5.8 percent) and non-Hispanic-owned (8.2 percent) businesses in the educational services

sector; and for male- (5.1 percent) and non-Hispanic-owned (7.6 percent) businesses in the arts, entertainment, and recreation sectors. “Owned by individuals of other races” group includes American Indian and Alaska Native as well as Native Hawaiian and Other Pacific Islander. Businesses in these six sectors were most likely to experience adverse effects to their business operations as a result of the COVID-19 pandemic, according to our 2021 analysis of the U.S. Bureau of Labor Statistics’ 2020 Business Response Survey. Sector-level data on employer businesses are suppressed, indicated with (S), and not available to report. This impacts one sector owned equally by Hispanic and non-Hispanic individuals and two sectors owned by individuals of other races.

Table 13: Businesses Filing Employment Returns and Use of the Employee Retention Credit by Sex of Business Owner and Business Size, among the Study Population, 2020

Business size quartile	Sex of business owner	Percent among businesses filing employment tax returns	Percent among businesses using Employee Retention Credit (ERC)	Percentage point difference between ERC users and employment tax returns
Total across quartiles	Female	27.45%	30.54%	3.09
	Male	72.53	69.41	-3.12
	Unknown or missing sex	0.03	0.04	0.01
0 – 24th	Female	36.15	37.9	1.75
	Male	63.82	62.04	-1.78
	Unknown or missing sex	0.03	0.06	0.03
25th – 49th	Female	31.02	35.12	4.1
	Male	68.95	64.82	-4.14
	Unknown or missing sex	0.03	0.06	0.04
50th – 74th	Female	24.49	29.49	5
	Male	75.48	70.49	-4.99
	Unknown or missing sex	0.03	0.02	-0.01
75th – 100	Female	17.43	20.24	2.8
	Male	82.54	79.74	-2.8
	Unknown or missing sex	0.03	0.02	<-0.01

Source: GAO analysis of Internal Revenue Service taxpayer and Social Security Administration data. | GAO-22-104582

Note: Taxpayer data are as reported by taxpayers and subject to taxpayer reporting errors. Use of the Employee Retention Credit is as reported on employment tax forms 941, 943, and 944, and form 7200 for advance payment of the COVID-19 tax provisions. Our analysis included credits claimed on Schedule R and through amended returns associated with form 941. Business size is based on annual receipts reported on forms Schedule C, Schedule F, or 1120-S. Quarterly returns (forms 941 and 7200) from second through fourth quarter 2020 and annual returns for 2020, including electronically filed returns and paper filings, are included in the table. The study population included businesses filing employment tax returns, which were also either (1) sole proprietorships filing form 1040, Schedule C or F; or (2) S corporations with a single owner. Data are as of September 2021 to March 2022. Numbers may not add to 100 because of rounding. Results are not generalizable to the universe of businesses using the provision.

Table 14: Ratio of Percent of Female to Male-Owned Businesses Using Selected COVID-19 Tax Provisions by Business Size, among Eligible Businesses in the Study Population, 2020

	Ratio of percent of eligible female- to male-owned businesses using each provision			
Business size quartile	Employers: Paid sick and family leave credits	Employers: Payroll tax deferral	Self-employed: Paid sick and family leave credits	Self-employed: Payroll tax deferral
Total across quartiles	0.95	1.12	1.46	1.27
0 – 24th	1.1	1.18	1.56	1.29
25th – 49th	1.38	1.05	1.74	1.35
50th – 74th	1.6	1.2	1.6	1.3
75th – 100th	1.13	1.1	1.37	1.44

Source: GAO analysis of Internal Revenue Service taxpayer and Social Security Administration data. | GAO-22-104582.

Note: Taxpayer data are as reported by taxpayers and subject to taxpayer reporting errors. Use of the employer provisions is as reported on employment tax forms 941, 943, and 944, and form 7200 for advance payment of the COVID-19 tax provisions. Our analysis included credits claimed on Schedule R and through amended returns associated with form 941. Use of the self-employed provisions is as reported on form 1040, Schedule 3. Business size is based on annual receipts reported on forms Schedule C, Schedule F, or 1120-S. Quarterly returns (forms 941 and 7200) from second through fourth quarter 2020 and annual returns for 2020, including electronically filed returns and paper filings, are included in the table. For the employer provisions, the study population included businesses filing employment tax returns and which were also either (1) sole proprietorships filing form 1040, Schedule C or F; or (2) S corporations with a single owner. For the self-employed provisions, the study population included sole proprietorships filing form 1040, Schedule C or F, and which also filed a form 1040, Schedule SE. Data are as of September 2021 to March 2022. Results are not generalizable to the universe of businesses using the provisions. Female-owned businesses were less likely to use the employer leave credit than male-owned businesses across all quartiles, but were more likely to use the provision in each of the quartiles individually. This result occurred because the association between provision use and sex of the business owner is qualitatively different from the association between those same two variables holding constant business size.

Table 15: Businesses Filing Employment Tax Returns and Use of Employee Retention Credit by Estimated Race and Ethnicity of Owner and Size, among the Study Population, 2020

Business size quartile	Estimated race or ethnicity of business owner	Estimated percent among businesses filing employment tax returns	Estimated percent among businesses using Employee Retention Credit (ERC)	Estimated percentage point difference between ERC users and employment tax return filers
Total across quartiles	Asian, non-Hispanic	9.84%	12.92%	3.08
	Black or African American, non-Hispanic	6.14	5.68	-0.46
	Hispanic, any race	10.2	8.97	-1.22
	Other races, non-Hispanic	2.12	2.06	-0.06
	White, non-Hispanic	71.71	70.36	-1.34
0 – 24th	Asian, non-Hispanic	9.19	11.86	2.68
	Black or African American, non-Hispanic	7.73	7.53	-0.2

Business size quartile	Estimated race or ethnicity of business owner	Estimated percent among businesses filing employment tax returns	Estimated percent among businesses using Employee Retention Credit (ERC)	Estimated percentage point difference between ERC users and employment tax return filers
	Hispanic, any race	12.15	11.33	-0.82
	Other races, non-Hispanic	2.13	2.13	0
	White, non-Hispanic	68.8	67.15	-1.66
25th – 49th	Asian, non-Hispanic	9.88	14.3	4.42%
	Black or African American, non-Hispanic	6.18	5.66	-0.5
	Hispanic, any race	10.32	8.99	-1.34
	Other races, non-Hispanic	2.13	2.06	-0.07
	White, non-Hispanic	71.49	68.99	-2.49
50th – 74th	Asian, non-Hispanic	10.33	14.54	4.21
	Black or African American, non-Hispanic	5.47	4.72	-0.75
	Hispanic, any race	9.65	8.47	-1.18
	Other races, non-Hispanic	2.12	2.06	-0.06
	White, non-Hispanic	72.44	70.22	-2.23
75th – 100th	Asian, non-Hispanic	10	11.18	1.18
	Black or African American, non-Hispanic	5.07	4.81	-0.25
	Hispanic, any race	8.51	7.15	-1.36
	Other races, non-Hispanic	2.09	2.01	-0.09
	White, non-Hispanic	74.33	74.86	0.53

Source: GAO analysis of Internal Revenue Service taxpayer, Social Security Administration, and U.S. Census Bureau data. | GAO-22-104582

Note: Taxpayer data are as reported by taxpayers and subject to taxpayer reporting errors. Use of the Employee Retention Credit is as reported on employment tax forms 941, 943, and 944, and form 7200 for advance payment of the COVID-19 tax provisions. Our analysis included credits claimed on Schedule R and through amended returns associated with form 941. Business size was based on annual receipts reported on forms Schedule C, Schedule F, or 1120-S. Quarterly returns (forms 941 and 7200) from second through fourth quarter 2020 and annual returns for 2020, including electronically filed returns and paper filings, are included in the table. The study population included businesses filing employment tax returns and which were also either (1) sole proprietorships filing form 1040, Schedule C or F; or (2) S corporations with a single owner. Data are as of September 2021 to March 2022. Results are not generalizable to the universe of businesses using the provision. “Asian” includes Native Hawaiian and Pacific Islander, non-Hispanic individuals. “Other races” include non-Hispanic American Indian or Alaskan Natives, individuals of two or more races, and individuals in smaller racial groups. Numbers may not add to 100 because of rounding.

Appendix IV: Articles Included in Literature Review

We reviewed the listed studies to identify analytical methods for assigning race, ethnicity, and sex when it is missing from a dataset:

1. Adjaye-Gbewonyo, Dzifa; Robert A. Bednarczyk; Robert L. Davis; Saad B. Omer. “Using the Bayesian Improved Surname Geocoding Method (BISG) to Create a Working Classification of Race and Ethnicity in a Diverse Managed Care Population: A Validation Study.” *Health Services Research*, vol. 49, no. 1 (2014): 268-283.
2. Akee, Randall; Maggie R. Jones; Sonya R. Porter. “Race Matters: Income Shares, Income Inequality, and Income Mobility for All U.S. Races.” *Demography*, vol. 56, no. 3 (2019): 999-1021.
3. Anyon, Yolanda; Duan Zhang; Cynthia Hazel. “Race, Exclusionary Discipline, and Connectedness to Adults in Secondary Schools.” *American Journal of Community Psychology*, vol. 57 (2016): 342-352.
4. Bayer, Patrick; Marcus Casey; Fernando Ferreira; Robert McMillan. “Racial and Ethnic Price Differentials in the Housing Market.” *Journal of Urban Economics*, vol. 102 (2017): 91-105.
5. Bosqui, Tania J.; Aideen Maguire; Anne Kouvonen; David Wright; Michael Donnelly; Dermot O’Reilly. “Ethnic Density and Risk of Mental Ill Health – The Case of Religious Sectarianism in Northern Ireland: A Population Data Linkage Study.” *Health & Place*, vol. 47 (2017): 29-35.
6. Chiu, Maria; Michael Lebenbaum; Alice M. Newman; Juveria Zaheer; Paul Kurdyak. “Ethnic Differences in Mental Illness Severity: A Population-Based Study of Chinese and South Asian Patients in Ontario, Canada.” *Journal of Clinical Psychiatry*, vol. 77, no. 9 (2016): e1108-e1116.
7. DeFilippis, Ersilia M.; Lauren Sinnenberg; Nadim Mahmud; Malissa J. Wood; Sharonne N. Hayes; Erin D. Michos; Nosheen Reza. “Gender Differences in Publication Authorship During COVID-19: A Bibliometric Analysis of High-Impact Cardiology Journals.” *Journal of the American Heart Association*, vol. 10, no. 5 (2021): 1-6.
8. Derose, Stephen F.; Richard Contreras; Karen J. Coleman; Corinna Koebnick; Steven J. Jacobsen. “Race and Ethnicity Data Quality and Imputation Using U.S. Census Data in an Integrated Health System: The Kaiser Permanente Southern California Experience.” *Medical Care Research and Review,* vol. 70, no. 3 (2012): 330-345.
9. Dhamoon, Mandip S.; Limei Zhou; Melissa Stamplecoski; Moira Kapral; Baiju Shah. “Stroke Recurrence among South Asians with

Diabetes in Ontario, Canada." *International Journal of Stroke*, vol. 11, no. 8 (2016): 890-897.

10. Elliott, Marc N.; Kirsten Becker; Megan K. Beckett; Katrin Hambarsoomian; Philip Pantoja; Benjamin Karney. "Using Indirect Estimates Based on Name and Census Tract to Improve the Efficiency of Sampling Matched Ethnic Couples from Marriage License Data." *Public Opinion Quarterly*, vol. 77, no. 1 (Spring 2013): 375-384.
11. Ennis, Sharon R.; Sonya R. Porter; James M. Noon; Ellen Zapata. "When Race and Hispanic Origin Reporting Are Discrepant across Administrative Records and Third Party Sources: Exploring Methods to Assign Responses." *Statistical Journal of the IAOS*, vol. 34, no. 2 (2018): 179-189.
12. Gerardi, Kristopher; Paul Willen; David Hao Zhang. "Mortgage Prepayment, Race, and Monetary Policy." *Working Paper Series (Federal Reserve Bank of Boston),* no. 20-7 (2020): 1-17.
13. Grundmeier, Robert W.; Lihai Song; Mark J. Ramos; Alexander G. Fiks; Marc N. Elliott; Allen Fremont; Wilson Pace; Richard C. Wasserman; and Russell Localio. "Imputing Missing Race/Ethnicity in Pediatric Electronic Health Records: Reducing Bias with Use of U.S. Census Location and Surname Data." *Health Services Research*, vol. 50, no. 4 (2015): 946-960.
14. Haas, Ann; Marc N. Elliott; Jacob W. Dembosky; John L. Adams; Shondelle M. Wilson-Frederick; Joshua S. Mallett; Sarah Gaillot; Samuel C. Haffer; Amelia M. Haviland. "Imputation of Race/Ethnicity to Enable Measurement of HEDIS Performance by Race/Ethnicity." *Health Services Research,* vol. 54 (2019): 13-23.
15. Harris, J. Andrew. "What's in a Name? A Method for Extracting Information about Ethnicity from Names." *Political Analysis*, vol. 23, no. 2 (Spring 2015): 212-224.
16. Hennessy, Deirdre A.; Andrea Soo; Daniel J. Niven; Rachel J. Jolley; Juan Posadas-Calleja; Henry T. Stelfox; Christopher J. Doig. "Socio-demographic Characteristics Associated with Hospitalization for Sepsis among Adults in Canada: a Census-linked Cohort Study." *Canadian Journal of Anesthesia*, vol. 67, no. 4 (2020): 408-420.
17. Hofstra, Bas and Niek C de Schipper. "Predicting Ethnicity with First Names in Online Social Media Networks." *Big Data & Society*, (January-June 2018): 1-14.
18. Howell, Sabrina T.; Theresa Kuchler; David Snitkof; Johannes Stroebel; Jun Wong. "Racial Disparities in Access to Small Business

Credit: Evidence from the Paycheck Protection Program." *National Bureau of Economic Research Working Paper*, no. 29364 (October 2021).

19. Imai, Kosuke and Kabir Khanna. "Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Records" *Political Analysis*, vol. 24, no. 2 (Spring 2016): 263-272.
20. Labgold, Katie; Sarah Hamid; Sarita Shah; Neel R. Gandhi; Allison Chamberlain; Fazle Khan; Shamimul Khan; Sasha Smith; Steve Williams; Timothy L. Lash; Lindsay J. Collin. "Estimating the Unknown: Greater Racial and Ethnic Disparities in COVID-19 Burden after Accounting for Missing Race and Ethnicity Data." *Epidemiology*, vol. 32 (2021): 157-161.
21. Meriküll, Jaanika; Merike Kukk; Tairi Rõõm. "What Explains the Gender Gap in Wealth? Evidence from Administrative Data." *Review of Economics of the Household*, vol.19, no. 2 (2021): 501-547.
22. Morgan, Steven; Gillian Hanley; Colleen Cunningham; Hude Quan. "Ethnic Differences in the Use of Prescription Drugs: A Cross-Sectional Analysis of Linked Survey and Administrative Data." *Open Medicine*, vol. 5, no. 2 (2011): e87-e93.
23. Olson, Marin; Rebecca J. Shlafer; Peter Bodurtha; Jonathan Watkins; Courtney Hougham; Tyler N. A Winkelman. "Health Profiles and Racial Disparities among Individuals on Probation in Hennepin County, Minnesota, 2016: A Cross-Sectional Study." *BMJ Open*, vol. 11, no. 9 (2021): e047930.
24. Pennap, Dinci; Mehmet Burcu; Daniel J. Safer; Julie M. Zito. "Hispanic Residential Isolation, ADHD Diagnosis and Stimulant Treatment among Medicaid-Insured Youth." *Ethnicity & Disease*, vol. 27, no. 2 (Spring 2017): 85-94.
25. Pico, T.; P. Bierman; K. Doyle; S. Richardson. "First Authorship Gender Gap in the Geosciences." *Earth and Space Science*, vol. 7 (2020): e2020EA001203.
26. Pine, Michael; Niranjana M. Kowlessar, Jason L. Salemi, Jill Miyamura; David S. Zingmond; Nicole E Katz; Joe Schindler. "Enhancing Clinical Content and Race/Ethnicity Data in Statewide Hospital Administrative Databases: Obstacles Encountered, Strategies Adopted, and Lessons Learned." *Health Services Research*, 50:S1, Part II (August 2015): 1300-1321.
27. Ryan, Ronan; Sally Vernon; Gill Lawrence; Sue Wilson. "Use of Name Recognition Software, Census Data and Multiple Imputation to Predict Missing Data on Ethnicity: Application to Cancer Registry Records."

BMC Medical Informatics and Decision Making, vol. 12, no. 3 (2012): 1-8.

28. Santamaría, Lucia and Helena Mihaljevic. “Comparison and Benchmark of Name-To-Gender Inference Services.” *PeerJ Computer Science*, (2018): 4:e156.
29. Silva, Gabriella C.; Amal N. Trivedi; Roee Gutman. “Developing and Evaluating Methods to Impute Race/Ethnicity in an Incomplete Dataset.” *Health Services and Outcomes Research Methodology*, vol. 19 (2019): 175-195.
30. Smith, Caroline K.; David K. Bonauto. “Improving Occupational Health Disparity Research: Testing a Method to Estimate Race and Ethnicity in a Working Population.” *American Journal of Industrial Medicine*, vol. 61, no. 8 (August 2018): 640-648.
31. So, Lawrence; Stephen G. Morgan: Hude Quan. “Does Concordance between Survey Responses and Administrative Records Differ by Ethnicity for Prescription Medication?” *Journal of Population Therapeutics and Clinical Pharmacology*, vol.19, no. 2 (2012): e248-e258.
32. Storey, Philip; Ann P. Murchison; Yang Dai; Lisa Hark; Laura T. Pizzi; Benjamin E. Leiby; Julia A. Haller. “Comparing Methodologies for Imputing Ethnicity in an Urban Ophthalmology Clinic.” *Ophthalmic Epidemiology*, vol. 21, no. 2 (2014): 106-110.
33. Szymkowiak, Marysia and Melissa Rhodes-Reese. “Addressing the Gender Gap: Using Quantitative and Qualitative Methods to Illuminate Women’s Fisheries Participation.” *Frontiers in Marine Science*, vol. 7, no. 299 (May 2020).
34. Thomas, Emma G.; Bamini Jayabalasingham; Tom Collins; Jeroen Geertzen; Chinh Bui; Francesca Dominici. “Gender Disparities in Invited Commentary Authorship in 2459 Medical Journals.” *JAMA Network Open*, vol. 2, no. 10 (2019).
35. Timbie, Justin W.; Ashley M. Kranz; Maria DeYoreo; Blen Eshete-Roesler; Marc N. Elliott; José J. Escarce; Mark E. Totten; Cheryl L. Damberg. “Racial and Ethnic Disparities in Care for Health System-Affiliated Physician Organizations and Non-Affiliated Physician Organizations.” *Health Services Research*, vol. 55, suppl. 3 (2020): 1107-1117.
36. Voicu, Ioan. “Using First Name Information to Improve Race and Ethnicity Classification.” *Statistics and Public Policy*, vol. 5, no. 1 (2018): 1-13.

37. Wais, Kamil. “Gender Prediction Methods Based on First Names with genderizeR.” *The R Journal*, vol. 8, no. 1 (August 2016): 17-37.

38. Xue, Yishu; Ofer Harel; Robert H. Aseltine, Jr. “Imputing Race and Ethnic Information in Administrative Health Data.” *Health Services Research*, vol. 54 (2019): 957-963.

39. Zuberi, Anita and Samantha Teixeira. “Death and Taxes: Examining the Racial Inequality in Premature Death across Neighborhoods.” *Journal of Community Psychology*, vol.49, no. 7 (2021): 2348-2365.

We reviewed the listed studies to understand the causes of low use of tax provisions as described in our interviews, and to identify additional factors that may affect use.

1. Bhargava, Saurabh and Dayanand Manoli. “Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment.” *American Economic Review*, vol. 105, no. 11(2015): 3489-3529.

2. GAO. *Small Employer Health Tax Credit: Factors Contributing to Low Use and Complexity.* GAO-12-549. Washington, D.C.: May 14, 2012.

3. Goldin, Jacob. “Tax Benefit Complexity and Take-up: Lessons from the Earned Income Tax Credit.” *Tax Law Review*, vol.72, no. 1 (2018): 59-110.

4. Goodman, Lucas. “Take-up of Payroll Tax-Based Subsidies during the COVID-19 Pandemic.” Working paper prepared by staff of the U.S. Department of the Treasury, Office of Tax Analysis. November 2021.

5. Kitchen, John and Matthew Knittel. “Business Use of Section 179 Expensing and Bonus Depreciation, 2002-2014.” Working paper prepared by staff of the U.S. Department of the Treasury, Office of Tax Analysis. October 2016.

6. Zwick, Eric. “The Costs of Corporate Tax Complexity.” *American Economic Journal: Economic Policy*, vol. 13, no. 2 (2021): 467-500.

Appendix V: Comments from the Internal Revenue Service

OFFICE OF THE
CHIEF RISK OFFICER

DEPARTMENT OF THE TREASURY
INTERNAL REVENUE SERVICE
WASHINGTON, DC 20224

July 14, 2022

Ms. Jessica K. Lucas-Judy
Director, Tax Policy and Administration, Strategic Issues Team
U.S. Government Accountability Office
441 G Street, N.W.
Washington, D.C. 20548

Dear Ms. Lucas-Judy:

On behalf of the Commissioner and the Senior Leadership team at the Internal Revenue Service, thank you for the opportunity to review your draft report titled: *TAX EQUITY: Enhanced Evaluation Could Improve Outreach to Small Business Owners* (GAO-22-104582).

The unprecedented COVID-19 pandemic illustrates the significant role that the IRS plays in the overall health of our country. We were called upon to take on new responsibilities impacting almost every American during this national crisis while also fulfilling our routine responsibilities of tax administration. We appreciate the input and recommendation from GAO which we take into careful consideration as we continue to make improvements to ensure the IRS can serve the needs of the nation's taxpayers well into the future.

GAO made one recommendation directed to IRS in their report. IRS has provided comments in the enclosure.

We appreciate having the opportunity to provide this response to the GAO report. If you have any questions, please contact me at Mark.E.Pursley@irs.gov. Thank you.

Sincerely,

Mark E. Pursley

Digitally signed by Mark E. Pursley
Date: 2022.07.14 11:29:05 -04'00'

Mark E. Pursley
IRS Chief Risk Officer

Enclosure

Enclosure

GAO Recommendation and IRS Response to GAO Report 22-104582, entitled **"TAX EQUITY: Enhanced Evaluation Could Improve Outreach to Small Business Owners".**

RECOMMENDATION 1
The commissioner of the IRS should evaluate IRS's outreach efforts to very small businesses and owners with diverse backgrounds, using relevant and complete information, to inform future outreach.

COMMENTS:
The IRS agrees in principle with this recommendation.

It should be noted that this recommendation does not recognize the complexity associated with reaching diverse groups of small business owners based on demographic data because currently the IRS does not track or collect this information. However, we will collaborate with Treasury and inside the IRS on this area, which would also help refine future education and outreach.

Currently, IRS outreach efforts have been and continue to focus on improving service and compliance to diverse communities by making materials available in multiple languages and offering events to diverse groups of stakeholders such as schools, national small business groups, tax practitioner organizations, chambers of commerce and many others.

IRS outreach functions continue to coordinate with their respective partner groups to share outreach materials and social media content through partner bundles; part of this includes small business organizations with diverse populations and those that help underserved individuals.

The IRS continuously considers and evaluates stakeholder feedback that is received from a wide variety of sources including tax practitioners from around the country that can elevate taxpayer concerns and needs for the agency to address in real-time. Stakeholder feedback helps inform our approach to improve tax administration and compliance. Feedback received is utilized to help the IRS better understand communication needs and informs outreach efforts for future events so that small business audiences can understand tax obligations and their eligibility for programs and benefits.

Appendix VI: GAO Contact and Staff Acknowledgments

GAO Contact

Jessica Lucas-Judy, (202) 512-6806 or LucasJudyJ@gao.gov

Staff Acknowledgments

In addition to the contact named above, Sonya Phillips (Assistant Director), Shelby Kain (Analyst-in-Charge), Jieun Chang, Jacqueline Chapin, Denise Cook, Caitlin Cusati, Shelby Gullion, Suzanne Kaasa, Ed Nannenhorn, Rachel Schultz, Dylan Stagner, Jeff Tessin, Sonya Vartivarian, Peter Verchinski, and Alicia White made key contributions to this report.

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through our website. Each weekday afternoon, GAO posts on its website newly released reports, testimony, and correspondence. You can also subscribe to GAO's email updates to receive notification of newly posted products.

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, https://www.gao.gov/ordering.htm.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

Connect with GAO

Connect with GAO on Facebook, Flickr, Twitter, and YouTube.
Subscribe to our RSS Feeds or Email Updates. Listen to our Podcasts.
Visit GAO on the web at https://www.gao.gov.

To Report Fraud, Waste, and Abuse in Federal Programs

Contact FraudNet:

Website: https://www.gao.gov/about/what-gao-does/fraudnet

Automated answering system: (800) 424-5454 or (202) 512-7700

Congressional Relations

A. Nicole Clowers, Managing Director, ClowersA@gao.gov, (202) 512-4400, U.S. Government Accountability Office, 441 G Street NW, Room 7125, Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, DC 20548

Strategic Planning and External Liaison

Stephen J. Sanford, Managing Director, spel@gao.gov, (202) 512-4707
U.S. Government Accountability Office, 441 G Street NW, Room 7814,
Washington, DC 20548