The Characteristics of Authors Reflected in Copyright Registrations

A research note by the U.S. Copyright Office,
Office of the Chief Economist

June 2026

The Characteristics of Authors Reflected in Copyright Registrations

A research note by the U.S. Copyright Office, Office of the Chief Economist

June 2026

Project team:

This research note was prepared by the U.S. Copyright Office's Office of the Chief Economist. Contributors to this research note include Michael Palmedo, Brent Lutes, and Ryan Safner.

Summary

This research note, produced by the U.S. Copyright Office's Office of the Chief Economist, examines socioeconomic patterns of authors who register copyrights. It is based on copyright registrations submitted by individuals (as compared to corporate or other organizational claimants) between 2011 and 2024,[1] which have been aggregated by zip code, year, and type of creative work. Statistical analyses of these data produce the following key findings:

- More copyright registrations originate from areas where people have higher levels of educational attainment, earn higher incomes, are older, and work in the information and entertainment industries.
- More copyright registrations per person originate from urban areas than rural or suburban ones, suggesting certain productive efficiency gains from agglomeration (the tendency for economic activity to cluster geographically).
- The relationship between copyright registrations and socioeconomic characteristics varies substantially for different types of creative works.

Introduction

The U.S. Copyright Office is conducting research to learn more about the people and organizations that use the copyright system to protect their works. Data drawn from copyright registrations can improve understanding of how the copyright system is used by those engaged in creative activity.

This research note describes some of the socioeconomic patterns in the protection of works through registered copyrights. It uses addresses reported by claimants in copyright registration applications to identify areas with varying levels of registration activity and matches these areas to socioeconomic data published by the U.S. Census Bureau.

The data show that more copyright registrations tend to originate from areas with higher levels of educational attainment, higher incomes, older populations, and greater employment in the information and entertainment industries. The relative strength of each socioeconomic factor varies across types of creative works, and the observed relationships persist when controlling for population and other factors. However, the analysis cannot distinguish between differences in underlying creative activity and differences in the propensity to register copyrights.

---

1 The U.S. Copyright Office's full public dataset of approximately 22 million copyright registration records and 15 million other records from January 1, 1978, to June 27, 2025, is available on **copyright.gov**. Because the Census Bureau annual data goes through 2024, this research note compares data through 2024 as well.

Data and Methodology

Data directly reflecting the socioeconomic characteristics of authors who register copyrights is not available. For that reason, this analysis relies on a geographic pattern matching approach to extrapolate such information.

This approach uses statistical techniques to measure the relationship between the socioeconomic composition of a geographic area in a given year with copyright registration volumes (by type of work) originating from that area in the same year. Importantly, this methodology does not intend to identify the socioeconomic characteristics of any individual author. Rather, it is intended only to draw out high-level relationships between the factors of interest in large groups of authors and their collective copyright registration behavior. Moreover, the results of this methodology can neither confirm nor preclude causal relationships; they can only identify and measure correlative relationships.

Additionally, the methodology identifies how the characteristics of groups relate to their copyright registration behavior, but it cannot definitively identify which individuals within the group may be submitting more or fewer registrations. For example, the methodology reveals that areas with an older population produce more registrations than younger areas. This, in turn, suggests that older individuals may be more likely to register copyrights relative to younger individuals, but it does not establish this conclusively. It is possible that the young individuals who reside among older populations behave differently with respect to copyright registration than do young people who reside among younger populations.[2] In order to extrapolate the findings in this research note to individual-level behaviors, one must assume that sub-groups (such as high-income or older individuals) do not differ across geographies in relevant ways. This assumption is untested and may be more likely to hold for certain characteristics than others. The remainder of this section provides further details on the data and methodology employed.

Copyright Registration Counts

The findings in this research note are based on data taken from copyright registrations between 2011 and 2024 for the most commonly registered types of works.[3] The main unit of analysis is the count of registrations by zip code,[4] which is obtained from the address entered by claimants on their applications for copyright registrations.

---

2 For example, younger people living among an older population may be adults who live with their parents and thus have different opportunities for producing and registering creative works.

3 The most commonly registered types of works in the data are the following: art works, dramatic works, literary works, machine-readable works (including software), motion pictures, musical works, serials, sound recordings, and registrations that protect sound recordings as well as the underlying work in the same registration. Between 2011 and 2024, there were 3,226,037 registrations by individuals of these work types. Other types of works for which the Office receives relatively few registrations are not included in this analysis. A full description of the types of works registered can be found in *Library of Congress Copyright Data as Distributed in the MARC 21 Format*, available at **copyright.gov/economic-research**.

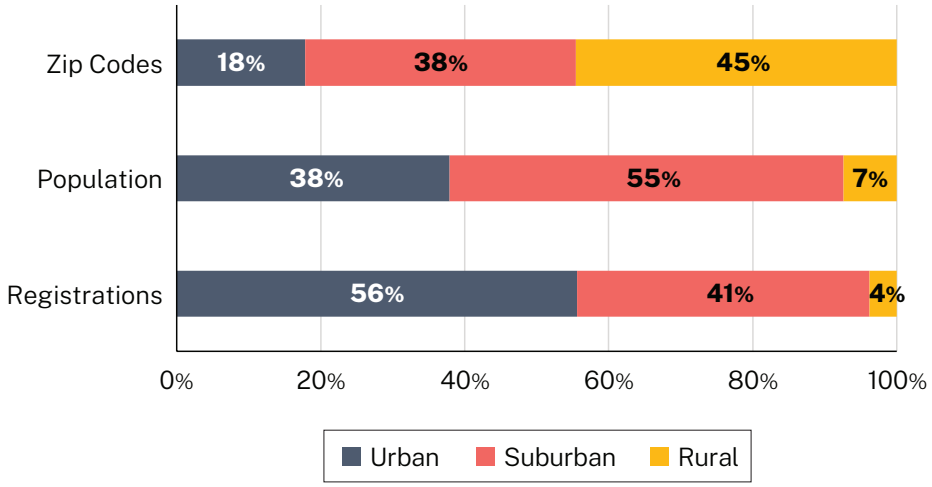
4 Zip codes do not represent geographic areas, so the report uses ZIP Code Tabulation Areas (ZCTAs), which provide geographic approximations of zip codes. See **census.gov/programs-surveys/geography/guidance/geo-areas/zctas.html**.

The analysis specifically draws data from registrations by individual claimants (which excludes registrations by firms or other organizations).[5] The dataset is restricted to registrations for which the first claimant listed on an application for a copyright registration is also listed as an author, which is true for 97.5 percent of the registrations by individuals during this time period.[6]

The dataset contains information on 27,813 zip codes. As illustrated by Figure 1, relatively few zip codes are urban, with rural zip codes accounting for the largest share. However, a larger share of the population resides in urban zip codes, and only a small share resides in rural ones. Just over a third of the zip codes in the dataset are suburban, but over half of the population resides in them.[7]

The majority of copyright registrations by individuals are registered by people who live in urban areas. As shown In Figure 1, only 38 percent of the U.S. population resides in zip codes that are classified urban by the Census Bureau, yet 56 percent of the copyright registrations by individuals originate from these zip codes. Conversely, a small percentage of Americans live in rural areas, and an even smaller share of copyright registrations originate from them. This indicates that the greater number of registrations from urban areas is not fully explained by their larger population.

Figure 1: Share of Zip Codes, Population, and Registrations in Dataset



5 The analysis in this research note is based on where an author resides. Addresses listed in their copyright registration applications are occasionally publishers or other third parties rather than the author's residence. However, in analyzing a random sample of these addresses, the vast majority (83 percent) proved to be residential addresses.

6 A majority of registrations by individuals contain one claimant and one author, but a substantial minority (27 percent) have more than one author. When multiple claimant addresses are provided, the methodology relies only on the address of the first claimant who is also an author.

7 This is based on data from the 2010 U.S. Decennial Census of Population and Housing, which reports the percentage blocks within each zip code that qualify as "urban" based on population and housing benchmarks. If the Census Bureau classifies a zip code as 100 percent urban, this report labels it urban; if the Census Bureau classifies the zip code as 0 percent urban, this report labels it rural; and if the Census Bureau classifies a zip code as between 0 and 100 percent urban, this report labels it suburban.

In line with this, the number of registrations remain higher in urban areas even when measured relative to population. Urban zip codes average 20.8 registrations per 100,000 people each year, whereas suburban and rural zip codes respectively average 7.6 and 9.7. To the degree that copyright registrations are reliable indicators of creative activity, this is consistent with economic literature on "agglomeration" effects of cities. This literature focuses on the benefits of geographic clustering of people and their latent knowledge, encouraging the flow of knowledge — both person-to-person and between people and firms — improving the overall productivity of the area when people and firms are closer together.[8] The phenomenon has been observed across the information industries, but it is thought to matter even more in the creative industries.[9]

Copyright registration counts are analyzed both in aggregate and separately by type of work. Work types are based on the MARC 21 bibliographic categories and include the following: artistic works, audio works,[10] dramatic works, literary works,[11] machine-readable works, and video works.[12]

Socioeconomic Variables

The U.S. Copyright Office's geographic data on copyright registrations is paired with socioeconomic data from the U.S. Census Bureau in a series of tests using different combinations of the variables listed below:

- Educational attainment, measured by the percentage of adults in each zip code with a bachelor's degree or higher degree.
- Income, measured by real median household income, in thousands of dollars.[13]
- Population density, measured as the total population in a zip code divided by its area.
- Median age, measured in years.
- Employment in relevant industries, specifically, the percentage of the workforce in a zip code that is employed in either the information sector or arts and entertainment sector.[14]

---

8 Edward L. Glaeser et al., "Cities, Skills and Regional Change," *Regional Studies* 48, no. 1 (2014): 7–43, DOI: 10.1080/00343404.2012.67463; Edward L. Glaeser et al., "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States," *Journal of Economic Literature* 47, no. 4 (2009): 983–1028, DOI: 10.1257/jel.47.4.983.

9 Ake Anderson et al., "Location and Spatial Clustering of Artists," *Regional Science and Urban Economics*. 47, no. C (2014): 128–137.

10 Several MARC 21 categories related to audio works — musical works, sound recordings, and registrations for both sound recordings and the underlying works — are combined into one category: "audio works." This includes music, as well as other types of audio works, such as podcasts and recorded sermons.

11 Two MARC 21 categories for written media — nondramatic literary works and serials — are combined into one category called "literary works."

12 The MARC 21 name for video works is "Motion Pictures and Film," but the works registered under this category include many non-feature-film-length videos, so this note uses the broader term "video works" to describe the full category.

13 Nominal dollar values reported by the U.S. Census Bureau were converted to real dollars using the U.S. Bureau of Labor Statistic's Consumer Price Index Retroactive Series. **bls.gov/cpi/research-series/r-cpi-u-rs-home.htm**.

14 Employment by sector is based on employment in NAICS 51 ("Information") or NAICS 71 ("Arts, Entertainment, and Recreation").

Statistical Tests

A series of statistical tests examines the relationship between the number of copyright registrations by individuals in each zip code and the socioeconomic factors described above. The results, presented in the following section, estimate the strength of the association between registration volume in a given zip code and year and each of these factors. Estimates are reported for the full set of registrations — including all types of works — and separately for each work type.

Because the analysis relies on count data, which are highly concentrated in more populous areas and include many zero observations, a method designed for count data is used to ensure that comparisons across areas are meaningful and not driven simply by differences in size.[15] This approach facilitates meaningful comparisons across localities of different sizes and levels of activity. The model also accounts for systematic differences across work types and for trends over time. The estimates can be interpreted as the expected change in registrations associated with a change in a given socioeconomic factor, holding other factors constant.

These tests build on previous geographic analysis by the Copyright Office, which identifies claimants' addresses[16] and patterns in registration activities.[17]

15 Specifically, the relationship is tested using Poisson regressions on registrations counts per zip code, year, and type of work. The independent variables are population and the socioeconomic factors described in the previous subsection. Total population, population density, and median age are log transformed. Fixed effects for the type of work and year are included. Errors are clustered by the type of work and year for the regressions on the full set, and by year for the regressions on individual types of works. Outliers representing 0.01 percent of the observations were dropped.

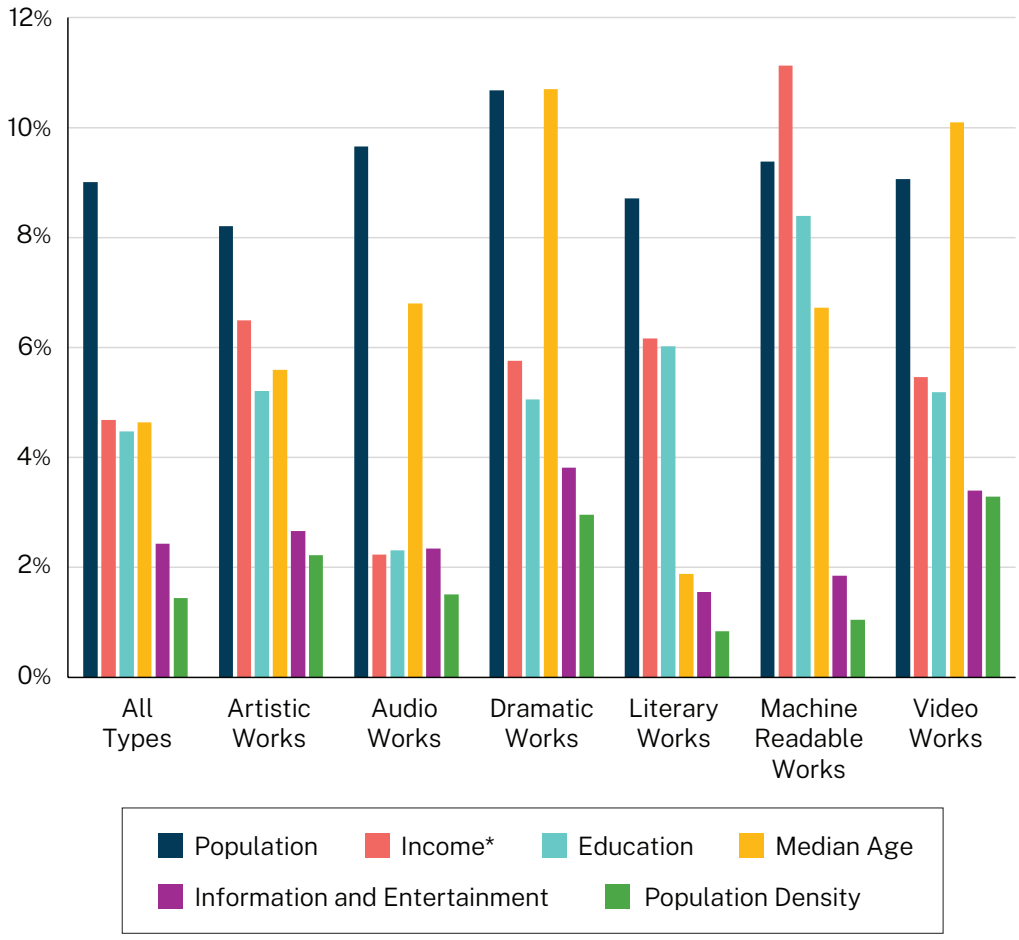
16 Brent Lutes et al., "Extending Intellectual Property Research in Copyright: A New Dataset from the U.S. Copyright Office," *Strategy Science* 10, no. 3 (2025): 245–262.

17 U.S. Copyright Office, *The Geography of Copyright Registrations*, (U.S. Copyright Office, 2024), **copyright.gov/economic-research/geography**.

Results

Figure 2 shows the relationship between registrations and six socioeconomic factors. Each bar reflects the percentage increase in copyright registrations that would be associated with a 10 percent increase in one of the factors, while holding other factors fixed.

Figure 2: Estimated Percentage Increases in Copyright Registrations Associated With a Ten Percent Rise in Each Socioeconomic Factor



* Income and education are highly correlated and therefore are not included in the same model specification. Two models are estimated: a baseline model that includes education (but not income) with all other factors, and an alternative model that replaces education with income. Figure 2 reports coefficients from the baseline model for all variables, along with the coefficient on income from the alternative model. Estimates for the other variables are similar across both specifications. Because income and education are highly correlated, the estimated effect of either variable should be interpreted as reflecting their combined association with registration.

The first set (of the six sets in Figure 2) of five bars reports the results for tests on aggregated data, which includes all work types. Population size is the factor most strongly correlated with copyright activity in an area. A 10 percent increase in the population of a zip code is associated with a 9 percent increase in copyright registrations (while holding all other socioeconomic factors constant), revealing a nearly one-for-one relationship between the two in percentage terms. Median household income, the share of residents with a college degree, and median age are also strongly correlated with registration activity. A 10 percent increase in each of these variables is associated with increases in expected registrations of 4.7 percent, 4.5 percent, and 4.6 percent, respectively. The share of workers employed in the information and entertainment sectors, as well as population density, also have positive associations with the number of copyright registrations, although the associations are not as strong.

The next five sets of bars in Figure 2 present results from models estimated separately for each type of work. Population exhibits the strongest relationship with registrations for artistic, audio, and literary works, while age is most strongly associated with registration activity for dramatic and video works. Income is most strongly associated with registrations of machine-readable works and is the second strongest factor for artistic and literary works. Education is also strongly associated with most work types: a 10 percent increase in the share of college graduates in a zip code is associated with more than a 5 percent increase in registrations for every category except audio works. Finally, both the share of workers in information and entertainment industries and population density are positively and statistically significantly associated with registrations, although their effects are comparatively modest.

The strength of these relationships varies across work types, particularly for education and income. A 10 percent increase in the share of the population with a college degree is associated with an 8.4 percent increase in registrations of machine-readable works, compared to 2.3 percent for audio works. Similarly, a 10 percent increase in median household income is associated with an 11.2 percent increase in registrations of machine-readable works, but only a 2.2 percent increase for audio works.

Conclusion

The analysis discussed in this research note contributes to a broader understanding of the socioeconomic patterns associated with copyright registration activity. The findings indicate that copyright registrations by individuals are more common in areas with higher levels of educational attainment, higher incomes, older populations, greater concentrations of workers in information and entertainment industries, and higher levels of urbanization, although the strength of these relationships varies across different types of creative works. Further research may help refine understanding of the relationship between socioeconomic conditions, geography, and participation in the copyright system.

UNITED STATES COPYRIGHT OFFICE
101 INDEPENDENCE AVE. SE
WASHINGTON, DC 20559-6000

COPYRIGHT.GOV

PUBLISHED JUNE 2026 / M-370