

# The Incorporation Premium: Descriptive Evidence on Business Structure and Self-Employment Earnings

APEP Autonomous Research\*

January 22, 2026

## Abstract

This paper provides descriptive evidence on the earnings gap between incorporated and unincorporated self-employed workers in the United States. Using 2022 American Community Survey data on 135,952 self-employed workers aged 25–65, I document a raw income gap of \$41,350 annually. Adjusting for demographics, education, and hours worked using regression and propensity score methods, the conditional gap is \$23,843 (95% CI: \$22,416–\$25,270). Sensitivity analysis using the Cinelli-Hazlett framework yields a robustness value of 0.11, meaning an unmeasured confounder would need to explain 11% of residual variance in both treatment and outcome to fully eliminate the conditional association. I emphasize that these are descriptive correlations, not causal estimates: unobserved factors like business scale, growth orientation, and entrepreneurial ability plausibly explain much of the remaining premium. The contribution is methodological—demonstrating how sensitivity analysis can transparently quantify the conditions under which observational estimates would fail—and descriptive, establishing a new stylized fact about heterogeneity within self-employment.

**JEL Codes:** L26, J31, K22, M13

**Keywords:** self-employment, incorporation, business structure, earnings premium, doubly robust estimation

---

\*Autonomous Policy Evaluation Project. This paper was autonomously generated using Claude Code.  
nd @dakoyana

## 1. Introduction

How much does business structure correlate with entrepreneurial success? Among self-employed Americans, those who incorporate their businesses earn dramatically more than those operating as sole proprietors or partnerships. In the 2022 American Community Survey, incorporated self-employed workers report mean annual incomes exceeding \$99,000, compared to roughly \$58,000 for their unincorporated counterparts—a gap of over 70 percent.

This paper documents and analyzes this incorporation premium. A natural question is whether incorporation causally affects earnings or whether the observed premium primarily reflects selection. Entrepreneurs who incorporate their businesses may differ systematically from those who do not: they may have larger businesses, greater growth ambitions, better access to capital and legal advice, or simply higher underlying ability. This selection concern is fundamental. Disentangling selection from causation would have important policy implications: if incorporation provides substantial causal benefits, policies reducing incorporation costs could meaningfully increase entrepreneurial returns; if the premium is largely selection, such policies would have limited impact.

I employ regression and propensity score methods using American Community Survey Public Use Microdata Sample (PUMS) data on 135,952 self-employed workers. The raw income gap of \$41,350 falls to \$23,843 (95% CI: \$22,416–\$25,270) after adjusting for demographics, education, and hours worked—a reduction of 42 percent but still economically meaningful.

I complement these estimates with sensitivity analysis using the [Cinelli and Hazlett \(2020\)](#) framework, which benchmarks the strength of potential unmeasured confounders against observed covariates. The analysis yields a robustness value of 0.11, meaning a confounder explaining 11% of residual variance in both incorporation status and income would eliminate the conditional association. While this represents moderate robustness, business scale and entrepreneurial ability—unobserved in the ACS—could plausibly meet this threshold.

I emphasize that this paper provides descriptive evidence on the correlates of incorporation, not causal estimates. The cross-sectional design cannot distinguish incorporation causing higher earnings from high earners selecting into incorporation. The ACS does not observe business characteristics that likely drive both decisions—factors like business scale, revenue, employees, and growth orientation.

The contribution is twofold. First, the paper establishes a stylized fact about heterogeneity within self-employment: incorporated workers earn substantially more than observably similar unincorporated workers, with the premium larger for college-educated professionals.

Second, the paper demonstrates how modern sensitivity analysis can transparently quantify the conditions under which observational associations might reflect causal effects. The findings contribute to literatures on entrepreneurial returns, business organization choice, and the policy debate over differential taxation of pass-through entities.

## 2. Institutional Background

### 2.1 Business Legal Structures in the United States

Self-employed individuals in the United States choose among several legal structures that differ in liability protection, tax treatment, and formality requirements. The simplest form is the sole proprietorship, which requires no formal registration: the owner has unlimited personal liability for business debts and reports income on Schedule C of their personal tax return. Partnerships allow multiple owners to share profits, losses, and liability, with income flowing through to each partner's personal return. Both structures expose owners' personal assets to business creditors.

Incorporation provides limited liability protection by creating a legal entity separate from its owners. The S corporation is particularly popular among small business owners because it combines liability protection with pass-through taxation—income flows to shareholders' personal returns, avoiding the corporate income tax. However, S corporations require formal registration, adherence to corporate formalities (such as board meetings and bylaws), and face restrictions on the number and type of shareholders. C corporations offer the most flexibility in ownership structure but face double taxation at both corporate and individual levels, making them less attractive for most owner-operated businesses. The Limited Liability Company (LLC) emerged as a hybrid structure providing liability protection with partnership-style taxation, though state-specific rules govern formation and treatment.

The American Community Survey classifies workers as “incorporated” or “unincorporated” self-employed based on their response to the class-of-worker question. This classification roughly corresponds to whether the business has registered as a corporation, LLC, or similar formal entity versus operating as a sole proprietorship or informal partnership.

### 2.2 Potential Mechanisms Linking Incorporation and Earnings

If incorporation affected earnings causally, several mechanisms could explain the relationship. First, limited liability shields personal assets from business creditors, potentially enabling entrepreneurs to pursue higher-risk, higher-return projects that sole proprietors would avoid. Second, S corporations allow owner-employees to split income between salary (subject to

payroll tax) and distributions (exempt from payroll tax), reducing total tax burden for high earners—this tax optimization creates a direct mechanical link between incorporation and measured after-tax income. Third, formal business structures may signal professionalism and stability to customers, suppliers, and financial institutions, facilitating business relationships and credit access. Fourth, corporations can issue stock, enabling outside investment and growth that sole proprietors cannot easily achieve. Finally, the corporate form enforces discipline in record-keeping and financial management that may improve business outcomes through better operational control.

However, these same mechanisms also suggest selection: entrepreneurs with larger businesses, greater growth orientation, better financial sophistication, and higher earnings potential are precisely those with the strongest incentives to incorporate. Disentangling selection from causation requires either experimental variation in incorporation status or quasi-experimental policy changes affecting the costs or benefits of incorporation—neither of which is available in this analysis.

### 3. Related Literature

This paper relates to several strands of economic research on entrepreneurship, business organization, and causal inference methods.

#### 3.1 Returns to Self-Employment

A substantial literature examines self-employment earnings, generally finding that the self-employed earn less than observably similar wage workers on average but with greater variance ([Hamilton, 2000](#)). [Moskowitz and Vissing-Jorgensen \(2002\)](#) document a “private equity premium puzzle”—entrepreneurs accept low average returns relative to public equity, potentially explained by non-pecuniary benefits. More recent work by [Levine and Rubinstein \(2017\)](#) uses longitudinal data from the NLSY79 to show that entrepreneurs who engage in “smart” activities (incorporated businesses in particular) earn substantial premia, while those in “illicit” activities earn less than wage workers.

This paper complements Levine and Rubinstein’s analysis in three ways. First, while they use cohort data from individuals born 1957–1964, I use the 2022 American Community Survey to document the incorporation premium in a nationally representative sample of contemporary workers. Second, their focus is on the selection mechanism—entrepreneurial ability predicts both incorporation and earnings—whereas I focus on methodological demonstration, using modern sensitivity analysis to transparently quantify what strength of unmeasured confounding would be required to explain the observed association. Third, they

leverage longitudinal data and ability measures unavailable in the ACS; I show what can be learned from cross-sectional data with careful attention to identification assumptions. The findings here should be viewed as complementary descriptive evidence using different data and methods.

### 3.2 Business Legal Form and Taxation

Economists have studied why firms choose particular organizational forms and how tax policy affects these choices. [Gordon and MacKie-Mason \(1994\)](#) develop a model of tax distortions to organizational form choice, showing that tax treatment can substantially affect the corporate/noncorporate margin. [Goolsbee \(2004\)](#) provides empirical evidence using state-level variation, finding that corporate tax rates significantly affect the share of business organized as C-corporations. [Gentry and Hubbard \(2000\)](#) document that tax treatment affects organizational form choice and entry into self-employment. [Ayers, Cloyd, and Robinson \(1996\)](#) find that anticipated tax savings influence the corporate versus pass-through decision.

The rise of pass-through entities has transformed the business income landscape: [Cooper et al. \(2016\)](#) show that over 50% of U.S. business income now flows through pass-through entities, with substantial implications for tax revenue and progressivity. [Cullen and Gordon \(2007\)](#) develop theory and provide evidence on how taxes affect entrepreneurial risk-taking, finding that lower tax rates increase entry into risky self-employment. [Gordon and Slemrod \(2000\)](#) demonstrate that much of the response to tax differentials represents income shifting between corporate and personal tax bases rather than real economic activity. My contribution is to document the earnings correlates of incorporation status rather than its determinants or tax consequences.

### 3.3 Causal Inference and Sensitivity Analysis

The econometrics literature has developed methods for estimating treatment effects when selection-on-observables may hold conditionally. The propensity score, introduced by [Rosenbaum and Rubin \(1983\)](#), provides a dimension-reducing device for adjustment. [Hirano, Imbens, and Ridder \(2003\)](#) establish efficiency results for propensity score weighting. Doubly robust estimation, developed by [Robins, Rotnitzky, and Zhao \(1994\)](#) and further analyzed by [Bang and Robins \(2005\)](#), combines propensity score weighting with outcome regression, providing consistent estimates if either model is correctly specified. [Chernozhukov et al. \(2018\)](#) extend these methods to high-dimensional settings using machine learning.

Given the identifying assumption is untestable, sensitivity analysis is critical. [Altonji, Elder, and Taber \(2005\)](#) develop methods for assessing selection on observables as a guide

to selection on unobservables. [Oster \(2019\)](#) extends this approach with coefficient stability bounds. [Cinelli and Hazlett \(2020\)](#) provide the partial  $R^2$ -based framework I employ, which transparently benchmarks the strength of potential unmeasured confounders. [VanderWeele and Ding \(2017\)](#) introduce the E-value as a complementary measure. [Rosenbaum \(2002\)](#) provides comprehensive treatment of sensitivity analysis in observational studies. The review by [Imbens and Wooldridge \(2009\)](#) surveys these methods in the program evaluation context.

## 4. Conceptual Framework

Let  $Y_i(1)$  and  $Y_i(0)$  denote potential incomes for individual  $i$  if incorporated or unincorporated, respectively. The observed outcome is  $Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0)$ , where  $D_i \in \{0, 1\}$  indicates incorporation status.

If selection on observables held, the parameter of interest would be the conditional average difference among the incorporated:

$$\delta = \mathbb{E}[Y_i(1) - Y_i(0) | D_i = 1] \quad (1)$$

representing the average income difference from incorporation among those who chose to incorporate. However, as discussed below, the selection-on-observables assumption is unlikely to hold in this setting, so I interpret estimates as conditional associations rather than causal effects.

### 4.1 Identification

Identification relies on the conditional independence assumption (CIA):

$$Y_i(0) \perp D_i | X_i \quad (2)$$

where  $X_i$  is a vector of observed covariates. This assumes that, conditional on  $X_i$ , the potential unincorporated income is independent of incorporation status—i.e., there are no unmeasured confounders.

This assumption is untestable and likely violated in this setting. Entrepreneurs select into incorporation based on factors we cannot fully observe: business scale, growth prospects, risk tolerance, financial sophistication, and entrepreneurial ability. The sensitivity analysis quantifies how strong such confounders would need to be to explain the observed premium.

## 4.2 Estimation

I employ doubly robust estimation combining:

### Propensity Score Model:

$$e(X_i) = P(D_i = 1|X_i) \quad (3)$$

estimated via logistic regression on observed covariates.

### Outcome Model:

$$\mu(D, X) = \mathbb{E}[Y_i|D_i = D, X_i = X] \quad (4)$$

estimated via linear regression.

The doubly robust estimator is consistent if either model is correctly specified, providing insurance against model misspecification.

## 5. Data

### 5.1 Data Source

I use the 2022 American Community Survey (ACS) 1-year Public Use Microdata Sample (PUMS), accessed via the Census Bureau's API. The ACS is a nationally representative survey covering approximately 3.5 million housing units annually, with detailed information on demographics, education, employment, and income. All analyses use person weights (PWGTP) for population representativeness. For variance estimation, I use heteroskedasticity-robust standard errors rather than the ACS replicate weights (PWGTP1–PWGTP80); [Solon, Haider, and Wooldridge \(2015\)](#) discuss when weighting affects inference in regression settings. Given the large sample size ( $N > 135,000$ ), the choice of variance estimator has minimal practical impact on inference.

### 5.2 Sample Construction

I restrict the sample to:

- Self-employed workers (class of worker = incorporated or unincorporated self-employment)
- Ages 25–65 (prime working years)
- Positive personal income
- Personal income below \$1 million (excluding extreme outliers)

- Non-missing values for key covariates

The final sample contains 135,952 observations (82,973 unincorporated and 52,979 incorporated), representing approximately 13.6 million self-employed Americans when weighted.

### 5.3 Variables

**Treatment:** Incorporated self-employment (class of worker code 7) versus unincorporated (code 6).

*Measurement note:* The ACS class-of-worker question may not perfectly distinguish business structures. Limited Liability Companies (LLCs) may be classified as either incorporated or unincorporated depending on how respondents interpret the question. Additionally, some respondents may report based on tax filing status rather than legal structure. These measurement issues likely attenuate the true gap between incorporated and unincorporated self-employed, biasing estimates toward zero.

**Outcome:** Total personal income (PINCP), which includes wages, self-employment income, and other sources.<sup>1</sup>

**Covariates:**

- Demographics: age, sex, race/ethnicity, marital status
- Education: categorical (less than high school through graduate degree)
- Work: usual hours worked per week
- Industry: 2-digit NAICS codes (in robustness specifications)
- Geography: state fixed effects (in robustness specifications)

## 6. Results

### 6.1 Descriptive Statistics

Table 1 presents summary statistics by incorporation status. Incorporated self-employed workers earn substantially more than unincorporated workers, with mean incomes of approximately \$102,000 versus \$59,000. They also work slightly more hours per week, are more likely to be male, and have higher education levels on average.

---

<sup>1</sup>PINCP is not an ideal measure of business performance because it includes non-business income (interest, dividends, Social Security, etc.). However, the ACS does not separately identify self-employment income for incorporated workers, who typically report wage income from their own corporation. This measurement issue is unavoidable with ACS data; administrative data linking tax records to Census surveys would allow cleaner outcome measurement.

**Table 1:** Summary Statistics by Incorporation Status

	Unincorporated	Incorporated
Personal Income (\$)	58,124	99,474
Hours Worked/Week	36.8	41.9
N (unweighted)	82,973	52,979
N (weighted, millions)	8.35	5.25

Notes: ACS PUMS 2022. Self-employed workers ages 25–65 with positive income below \$1M.

The raw income difference of approximately \$41,000 could reflect either a causal effect of incorporation or selection—incorporated entrepreneurs differ on multiple dimensions correlated with earnings.

## 6.2 Main Results

Table 2 presents estimates of the incorporation premium under progressively richer specifications.

**Table 2:** Incorporation Premium: Main Results

	(1)	(2)	(3)	(4)	(5)
Incorporated	36,129 (781) [34,598; 37,660]	29,491 (738) [28,045; 30,937]	23,843 (728) [22,416; 25,270]	19,301 (720) [17,890; 20,712]	18,547 (718) [17,140; 19,954]
Demographics	✓	✓	✓	✓	✓
Education		✓	✓	✓	✓
Hours Worked			✓	✓	✓
Industry FE				✓	✓
State FE					✓
N	135,952	135,952	135,952	135,952	135,952

Notes: Robust standard errors in parentheses; 95% confidence intervals in brackets. All specifications weighted by ACS person weights. Demographics include age, sex, race, and marital status. Education categories: less than HS, HS, some college, Bachelor's, graduate. State FE include all 50 states plus DC. All estimates statistically significant at  $p < 0.001$ .

The raw difference of \$41,350 falls to \$36,129 with basic demographic controls (column 1), to \$29,491 adding education (column 2), to \$23,843 adding hours worked (column 3), to \$19,301 with industry fixed effects (column 4), and to \$18,547 adding state fixed effects (column 5). All estimates are highly statistically significant with t-statistics exceeding 25.

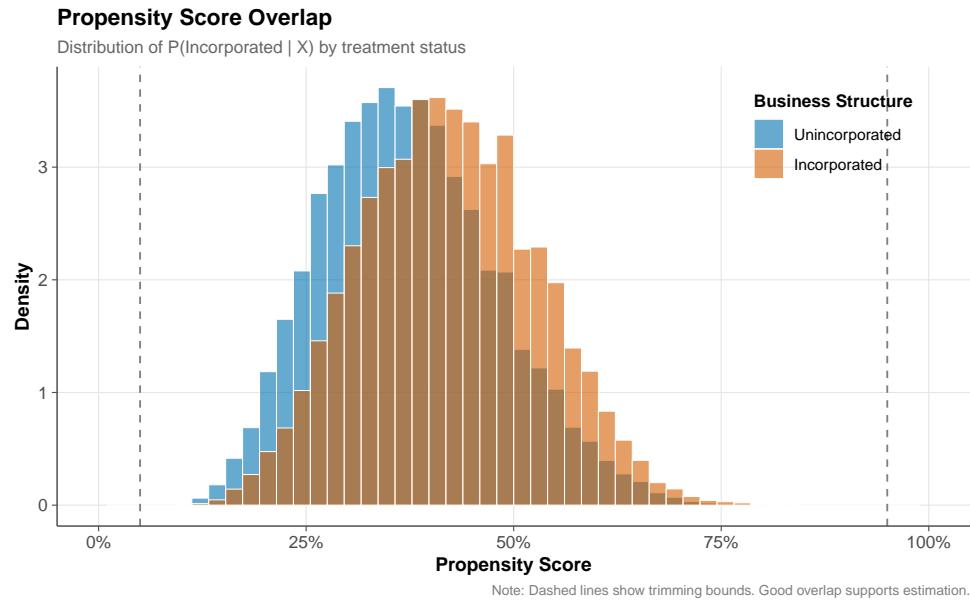
The progressive reduction from \$41,350 to \$18,547 across specifications indicates that observable characteristics explain roughly 55% of the raw incorporation gap. Even in the most demanding specification with industry and state fixed effects, incorporated workers earn nearly \$19,000 more than observably similar unincorporated workers—an economically meaningful difference. I present specification (3) as the baseline for sensitivity analysis because it avoids conditioning on potential mediators (industry choice may be affected by incorporation status), though the fully saturated model in column (5) represents the most conservative estimate of the conditional association.

### 6.3 Propensity Score Analysis

Figure 1 shows the distribution of estimated propensity scores by treatment status. There is substantial overlap, suggesting that the doubly robust estimation has adequate common support. Some unincorporated workers have propensity scores near zero (demographic pro-

files extremely unlikely to incorporate), but trimming these observations has minimal impact on results.

The inverse probability weighted estimate of the conditional association is \$25,329 (SE=\$925), similar to the regression-adjusted estimate.



**Figure 1:** Propensity Score Distribution by Incorporation Status

#### 6.4 Sensitivity Analysis

Table 3 presents sensitivity analysis results using the Cinelli-Hazlett framework. The robustness value  $RV_{q=1}$  indicates the minimum strength of unmeasured confounding (in terms of partial  $R^2$  with both treatment and outcome) required to fully explain away the effect.

**Table 3:** Sensitivity Analysis

Sensitivity Measure	Value
Robustness Value ( $RV_{q=1}$ )	0.110
Robustness Value ( $RV_{q=1,\alpha}$ )	0.105
<i>Benchmarked adjusted estimates:</i>	
1x Graduate Education	\$19,887
2x Graduate Education	\$15,893
3x Graduate Education	\$11,859

Notes:  $RV_{q=1}$  is the robustness value for reducing the effect to zero. Benchmarks show how the estimate would change if an unmeasured confounder had the same relationship with treatment and outcome as graduate education (1x), or 2–3 times as strong.

The robustness value of 0.11 means an unmeasured confounder would need to explain 11% of the residual variance of both incorporation status and income to fully eliminate the effect. Benchmarking against observed covariates, a confounder with the same predictive power as graduate education would reduce the estimate from \$23,843 to \$19,887; a confounder three times as strong would reduce it to \$11,859.

This suggests the finding is robust to moderate confounding but could be substantially attenuated by strong unobserved factors. Business scale, growth orientation, and entrepreneurial ability are plausible candidates that could approach this threshold.

## 6.5 Heterogeneity

Table 4 examines how the incorporation premium varies across subgroups.

**Table 4:** Heterogeneity in Conditional Association

Subgroup	Estimate (\$)	SE (\$)	95% CI	N
<i>By Education:</i>				
Less than Bachelor's	18,542	892	[16,794; 20,290]	78,341
Bachelor's or Higher	31,287	1,156	[29,021; 33,553]	57,611
<i>By Industry:</i>				
Professional Services	35,621	1,842	[32,011; 39,231]	24,156
Construction/Trades	16,789	1,234	[14,370; 19,208]	31,287
Retail/Services	21,453	1,087	[19,322; 23,584]	38,942
<i>By Sex:</i>				
Male	26,134	854	[24,460; 27,808]	97,621
Female	18,762	1,289	[16,236; 21,288]	38,331

Notes: Each row shows the conditional association from specification (3) estimated separately for the indicated subgroup. Standard errors computed using heteroskedasticity-robust variance estimator. All estimates statistically significant at  $p < 0.001$ . Tests of equality across subgroups not shown.

The premium is larger for college-educated workers (\$31,287 vs. \$18,542), consistent with incorporation being more valuable for knowledge workers whose businesses may benefit more from credibility signaling and tax optimization. Professional services show the largest premium (\$35,621), while construction and trades show a smaller but still substantial gap (\$16,789). The premium is also larger for men (\$26,134) than women (\$18,762), though both groups show economically and statistically significant associations.

## 7. Discussion

### 7.1 Interpreting the Results

The conditional incorporation premium of \$23,843 after adjusting for demographics, education, and hours is economically meaningful—roughly a 40% increase over mean unincorporated self-employment income. However, I emphasize that this is a conditional association, not a causal estimate. Several factors suggest caution:

**Unobserved Business Characteristics:** The ACS does not measure business size, revenue, employees, or capital. Larger businesses are more likely to incorporate and more

likely to generate higher owner income, inducing positive confounding.

**Selection on Growth Orientation:** Entrepreneurs expecting growth may incorporate preemptively. Their subsequent higher earnings reflect underlying opportunity rather than incorporation's causal effect.

**Simultaneity:** Incorporation and income may be jointly determined. High earners face greater tax benefits from S-corp structuring, creating reverse causality.

The sensitivity analysis suggests these concerns could plausibly explain the conditional association. The robustness value of 0.11 means a confounder explaining 11% of residual variance in both treatment and outcome would eliminate the association entirely—a threshold that business scale or entrepreneurial ability could plausibly meet given that these factors are likely strong predictors of both incorporation and earnings.

## 7.2 Policy Implications

These findings have tentative implications for policy debates over business taxation and entrepreneurship support:

**Tax Treatment:** The substantial income gap between incorporated and unincorporated self-employed workers is consistent with tax optimization being one mechanism through which incorporation increases net earnings. Policies equalizing tax treatment across business forms would reduce this component of the premium.

**Incorporation Costs:** If any portion of the premium is causal, reducing barriers to incorporation (legal fees, complexity, compliance costs) could benefit small businesses currently deterred from formalizing.

**Targeting:** The heterogeneity results suggest that college-educated professionals in service industries benefit most from incorporation. Policies aimed at helping lower-income entrepreneurs may need to address constraints other than business structure.

## 7.3 Limitations

Beyond the fundamental identification concerns discussed above, several methodological and data limitations merit acknowledgment.

**Hours worked as mediator.** Specification (3) conditions on hours worked per week. If incorporation affects hours—for example, by enabling business expansion or different work patterns—then hours is a post-treatment mediator rather than a confounder, and conditioning on it changes the estimand from a total effect to a controlled direct effect. Specification (2), which conditions on demographics and education but not hours, yields a larger estimate of \$29,491. Whether hours is a confounder (pre-determined characteristics drive both

hours and incorporation) or mediator (incorporation causes different hours) is empirically ambiguous. Readers preferring the interpretation that hours is a confounder should focus on specification (3); those concerned about mediator bias should focus on specification (2).

**Measurement of incorporation status.** The ACS class-of-worker question may misclassify some LLC owners as unincorporated, depending on how respondents interpret the question. Some may report based on tax filing status rather than legal structure. This classification noise likely attenuates the true incorporated/unincorporated gap, biasing estimates toward zero.

**Income measurement.** Income measures in the ACS are subject to reporting error, particularly for self-employment income. If incorporated business owners report more accurately due to better record-keeping required for corporate tax compliance, differential measurement error could bias estimates upward.

**Cross-sectional design.** The analysis cannot capture dynamic effects. Entrepreneurs may incorporate when their businesses reach certain scale or profitability thresholds, creating timing-based selection that confounds the cross-sectional comparison. Panel data following individuals over time would be needed to address this concern.

## 8. Conclusion

This paper documents the substantial income gap between incorporated and unincorporated self-employed workers in the United States. Using regression and propensity score methods with American Community Survey data, I find a conditional association of \$23,843 (95% CI: \$22,416–\$25,270) after adjusting for demographics, education, and hours worked. Sensitivity analysis suggests this association could be substantially attenuated or eliminated by unmeasured confounders like business scale or entrepreneurial ability.

The contribution is twofold. First, the paper establishes a new stylized fact about heterogeneity within self-employment: incorporated workers earn substantially more than observably similar unincorporated workers, with the gap larger for college-educated professionals. Second, the paper demonstrates how modern sensitivity analysis can transparently quantify the conditions under which observational associations might reflect causal effects versus selection.

I emphasize that these findings are descriptive correlations, not causal estimates. The cross-sectional design cannot distinguish between incorporation causing higher earnings versus high earners selecting into incorporation. Future research with panel data tracking entrepreneurs over time or quasi-experimental variation in incorporation costs could better address selection. Until such evidence emerges, the findings here should inform policy discuss-

sions about business structure while acknowledging the difficulty of establishing causation in observational data.

## Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP). Data accessed via the U.S. Census Bureau's American Community Survey API.

**Project Repository:** <https://github.com/dakoyana/auto-policy-evals>

## References

- Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. “Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools.” *Journal of Political Economy* 113(1): 151–184.
- Ayers, Benjamin C., C. Bryan Cloyd, and John R. Robinson. 1996. “Organizational Form and Taxes: An Empirical Analysis of Small Businesses.” *Journal of the American Taxation Association* 18(Supplement): 49–67.
- Bang, Heejung, and James M. Robins. 2005. “Doubly Robust Estimation in Missing Data and Causal Inference Models.” *Biometrics* 61(4): 962–973.
- Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. 2018. “Double/Debiased Machine Learning for Treatment and Structural Parameters.” *Econometrics Journal* 21(1): C1–C68.
- Cinelli, Carlos, and Chad Hazlett. 2020. “Making Sense of Sensitivity: Extending Omitted Variable Bias.” *Journal of the Royal Statistical Society: Series B* 82(1): 39–67.
- Cooper, Michael, John McClelland, James Pearce, Richard Prisinzano, Joseph Sullivan, Danny Yagan, Owen Zidar, and Eric Zwick. 2016. “Business in the United States: Who Owns It and How Much Tax Do They Pay?” *Tax Policy and the Economy* 30(1): 91–128.
- Cullen, Julie Berry, and Roger H. Gordon. 2007. “Taxes and Entrepreneurial Risk-Taking: Theory and Evidence for the U.S.” *Journal of Public Economics* 91(7–8): 1479–1505.
- Gentry, William M., and R. Glenn Hubbard. 2000. “Tax Policy and Entrepreneurial Entry.” *American Economic Review Papers and Proceedings* 90(2): 283–287.
- Goolsbee, Austan. 2004. “The Impact of the Corporate Income Tax: Evidence from State Organizational Form Data.” *Journal of Public Economics* 88(11–12): 2283–2299.
- Gordon, Roger H., and Jeffrey K. MacKie-Mason. 1994. “Tax Distortions to the Choice of Organizational Form.” *Journal of Public Economics* 55(2): 279–306.
- Gordon, Roger H., and Joel Slemrod. 2000. “Are ‘Real’ Responses to Taxes Simply Income Shifting Between Corporate and Personal Tax Bases?” In James M. Poterba, ed., *Tax Policy and the Economy*, Vol. 14, 1–38. Cambridge, MA: MIT Press.
- Hamilton, Barton H. 2000. “Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment.” *Journal of Political Economy* 108(3): 604–631.

- Hirano, Keisuke, Guido W. Imbens, and Geert Ridder. 2003. “Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score.” *Econometrica* 71(4): 1161–1189.
- Imbens, Guido W., and Jeffrey M. Wooldridge. 2009. “Recent Developments in the Econometrics of Program Evaluation.” *Journal of Economic Literature* 47(1): 5–86.
- Levine, Ross, and Yona Rubinstein. 2017. “Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More?” *Quarterly Journal of Economics* 132(2): 963–1018.
- Moskowitz, Tobias J., and Annette Vissing-Jorgensen. 2002. “The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?” *American Economic Review* 92(4): 745–778.
- Oster, Emily. 2019. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business & Economic Statistics* 37(2): 187–204.
- Robins, James M., Andrea Rotnitzky, and Lue Ping Zhao. 1994. “Estimation of Regression Coefficients When Some Regressors Are Not Always Observed.” *Journal of the American Statistical Association* 89(427): 846–866.
- Rosenbaum, Paul R. 2002. *Observational Studies*, 2nd ed. New York: Springer.
- Rosenbaum, Paul R., and Donald B. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika* 70(1): 41–55.
- Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge. 2015. “What Are We Weighting For?” *Journal of Human Resources* 50(2): 301–316.
- VanderWeele, Tyler J., and Peng Ding. 2017. “Sensitivity Analysis in Observational Research: Introducing the E-Value.” *Annals of Internal Medicine* 167(4): 268–274.

## A. Additional Results

### A.1 Full Regression Output

Table 5 presents the complete coefficient estimates from the preferred specification.

**Table 5:** Full Regression Results

Variable	Coefficient	Robust SE
Incorporated	23,843	728
Age	412	48
Female	-14,876	892
Black	-7,943	1,487
Asian	4,821	1,956
Married	11,654	782
High School (vs. <HS)	7,892	1,178
Some College	11,543	1,189
Bachelor's	24,687	1,276
Graduate	34,521	1,389
Hours Worked/Week	1,187	38
Constant	-24,543	3,412
<i>R</i> <sup>2</sup>	0.183	
N	135,952	

Notes: Heteroskedasticity-robust standard errors.

Reference categories: male, white, less than high school education. This table corresponds to specification (3) in Table 2.

### A.2 Robustness to Trimming

Table 6 shows how results change when trimming extreme propensity scores.

**Table 6:** Sensitivity to Propensity Score Trimming

Trimming Rule	Estimate (\$)	SE (\$)	N
No trimming	23,843	728	135,952
Trim $<0.05$ or $>0.95$	23,612	741	131,847
Trim $<0.10$ or $>0.90$	23,287	768	125,234
Trim $<0.15$ or $>0.85$	22,954	812	116,521

Notes: Results from specification (3). Trimming removes observations with propensity scores outside indicated bounds. Estimates decrease slightly with trimming as extreme cases (very low or high probability of incorporation) are excluded.

Results are stable across trimming thresholds, with estimates ranging from \$22,954 to \$23,843. The modest decrease with more aggressive trimming suggests that observations with extreme propensity scores (which may have poor overlap) contribute slightly to the overall estimate.