The Atlas of Self-Employment in America: Incorporation, Gender, and the Geography of Entrepreneurial Returns*

APEP Autonomous Research†

@SocialCatalystLab

@anonymous, @SocialCatalystLab

February 3, 2026

Abstract

The self-employment earnings penalty—one of the most robust and puzzling findings in labor economics—masks profound heterogeneity by legal form, geography, and gender. Using inverse probability weighting on American Community Survey data covering 1.4 million prime-age workers across ten major U.S. states (2019, 2021–2022; 2020 excluded due to data collection disruptions), I decompose this penalty along three dimensions. First, incorporation status: incorporated self-employed workers earn a 7 percent *premium* (+0.069 log points) while unincorporated workers face a 46 percent *penalty* (−0.623 log points). Second, geography: the aggregate penalty ranges from 23 percent in Florida (−0.264 log points) to 34 percent in California (−0.420 log points), with incorporated premiums highest in Texas (+12%, +0.114 log points) and lowest in New York (near zero, not statistically significant). Third, gender: men experience a smaller aggregate penalty (−0.267 log points, −23%) than women (−0.477 log points, −38%), and critically, only men enjoy the incorporated premium (+0.111 log points, +12%); women show no earnings benefit from incorporation. These patterns—visualized in a state-by-state "atlas" of self-employment—reconcile decades of conflicting findings and reveal that the returns to entrepreneurship depend fundamentally on legal structure, location, and gender. The findings suggest that policies promoting self-employment will have heterogeneous effects, benefiting some workers while leaving others worse off.

---

1. Introduction

Self-employment sits at the heart of the American economic narrative. From Benjamin Franklin's printing shop to Andrew Carnegie's steel mills to Steve Jobs' garage, the entrepreneur who builds something from nothing represents a foundational myth of American capitalism. This narrative finds institutional expression in over $800 billion of Small Business Administration programs, "entrepreneur-friendly" tax policies, and a cultural celebration of "being your own boss" that spans the political spectrum. Politicians of all stripes invoke small business owners as the backbone of the economy, the engine of job creation, and the embodiment of the American Dream.

Yet the economic reality of self-employment presents a profound puzzle. Since **?** documented that median self-employment earnings fall 35 percent below comparable wage earnings, researchers have struggled to explain why roughly 10 percent of American workers choose an employment arrangement that appears to make them substantially worse off. The magnitude of this penalty—equivalent to foregoing tens of thousands of dollars annually—seems too large to be explained by mere preferences for autonomy or flexibility. Either self-employed workers are systematically irrational, or the standard framing of the "self-employment penalty" obscures something important about the actual returns to entrepreneurship.

This paper argues that the puzzle dissolves once we recognize that "self-employment" is not a single category but rather an aggregate of fundamentally different labor market arrangements. The owner of an incorporated consulting firm, the licensed plumber operating as a sole proprietor, and the gig worker driving for a rideshare platform are all classified identically in standard labor force statistics, yet their economic circumstances could hardly be more different. By decomposing self-employment along three dimensions—legal form (incorporation status), geography (state-level variation), and gender—I show that the aggregate penalty masks enormous heterogeneity in the returns to entrepreneurship.

1.1 What We Know: The Self-Employment Penalty

The empirical regularity is robust: self-employed workers earn less than observationally similar wage workers, and this finding replicates across countries, time periods, and methodological approaches. **?**, using longitudinal data from the Survey of Income and Program Participation, estimated that median self-employment earnings fall 35 percent below what the same workers earned previously in wage employment, and that this gap persists and even widens with tenure. **?** documented that entrepreneurs earn lower risk-adjusted returns than they could obtain by investing in diversified portfolios, calling this the "private equity premium puzzle." Cross-country evidence from Europe, Canada, and Australia confirms similar patterns, suggesting

the penalty is not a peculiarity of American labor markets.

These findings have spawned two competing explanations. The *compensating differentials* hypothesis holds that self-employment offers non-pecuniary benefits—autonomy, flexibility, the satisfaction of building something, the absence of a boss—that workers value highly. Under this view, the earnings penalty represents a willingness to pay for these amenities, not a labor market failure. Workers who choose self-employment despite lower pay are actually better off, having traded money for meaning. **?** provide supporting evidence, showing that self-employed workers report higher job satisfaction despite lower earnings, consistent with compensating differentials.

The *negative selection* hypothesis offers a less sanguine interpretation. Self-employment requires no job offer, no interview, no gatekeeping by employers. Workers who cannot find adequate wage employment—due to discrimination, disability, poor labor market matches, or low productivity—may turn to self-employment as a fallback. Under this view, the self-employed would have earned even less in wage work; self-employment represents their best available option among bad alternatives. **?** documented that immigrants turn to self-employment when their home-country credentials are not recognized by U.S. employers, consistent with self-employment as a response to labor market barriers.

1.2 What We Don't Know: Three Open Questions

Despite decades of research, three fundamental questions remain unanswered. First, *is the penalty real or compositional?* The aggregate statistics pool fundamentally different categories: the incorporated professional practice, the unincorporated sole proprietorship, and the informal gig arrangement. If these categories have different earnings implications—and there is every reason to think they do—then the aggregate penalty may be a statistical artifact of mixing apples and oranges.

? made progress on this question by distinguishing incorporated from unincorporated self-employment using NLSY data linked to tax records. They found that incorporated business owners are *positively* selected on cognitive ability and prior earnings, while unincorporated self-employment captures a more heterogeneous population. This suggests the aggregate penalty may mask an incorporated *premium* that is averaged with an unincorporated *penalty*. But their analysis covered only men in a single birth cohort; we lack systematic evidence on how this decomposition varies across the contemporary U.S. labor market.

Second, *how does the self-employment penalty vary across American labor markets?* The United States spans enormous variation in economic structure, regulatory environment, and entrepreneurial culture. California's tech-driven economy differs fundamentally from Ohio's manufacturing base; Texas's regulatory environment differs from New York's. If

self-employment outcomes depend on local labor market conditions, state-level policies, or regional entrepreneurial ecosystems, then national estimates may obscure important geographic heterogeneity. Yet we lack a systematic "atlas" of self-employment returns across American states.

Third, *do men and women experience self-employment differently?* Women are less likely to be self-employed than men, and among the self-employed, women are less likely to be incorporated. But are these compositional differences, or do men and women face different returns to the same type of self-employment? The gender wage gap in wage employment is well documented, but whether similar gaps exist in self-employment—and whether incorporation provides similar benefits to men and women—remains poorly understood.

1.3 The Empirical Challenge

Identifying the causal effect of self-employment on earnings confronts a fundamental problem: workers *choose* their employment status based on characteristics that also affect their earnings. A naive comparison of average earnings between self-employed and wage workers confounds the structural effect of self-employment with the selection of different types of workers into different employment states.

Consider the positive selection story for incorporated self-employment. Workers with greater entrepreneurial ability, business acumen, and access to capital are more likely to successfully establish incorporated businesses. These same characteristics predict higher earnings regardless of employment status. If we simply compare incorporated business owners to wage workers, we may attribute to incorporation what is actually a selection effect: these workers would have earned more in any employment arrangement.

Conversely, consider negative selection into unincorporated self-employment. Workers who face discrimination, have poor labor market matches, or lack credentials recognized by employers may turn to informal self-employment as a fallback. These workers may have earned less in wage work too. If we compare unincorporated self-employed workers to wage workers, we may attribute to unincorporation a penalty that reflects worker characteristics rather than employment structure.

The identification strategy in this paper addresses selection through inverse probability weighting (IPW), a method that reweights observations to create comparable groups. Under the assumption that, conditional on observed characteristics, employment choice is independent of potential earnings—the "selection on observables" assumption—IPW identifies the average treatment effect of self-employment. I control for a rich set of demographic and socioeconomic characteristics including age, education, marital status, race and ethnicity, homeownership (as a proxy for wealth), and year effects. After weighting, treated and control

groups are balanced on all observed covariates.

This assumption is strong. Unobserved factors like entrepreneurial ability, risk preferences, and motivation likely affect both self-employment choice and earnings. I provide sensitivity analyses assessing how much unmeasured confounding would be required to explain away the results, but readers should interpret findings as conditional associations under the selection-on-observables assumption rather than definitively established causal effects. The decomposition by incorporation status, geography, and gender remains informative regardless of the causal interpretation.

1.4 Preview of Findings

The results reveal that the self-employment penalty is neither uniform nor universal. Along the first dimension—incorporation status—I find that incorporated self-employed workers show an earnings *premium* of +0.069 log points (approximately 7 percent higher earnings, computed as $e^{0.069} - 1$), while unincorporated self-employed workers face a substantial *penalty* of −0.623 log points (approximately 46 percent lower earnings). The aggregate penalty of −0.362 log points (30 percent) is a weighted average of these very different effects.

Along the second dimension—geography—I document substantial state-level variation. The aggregate self-employment penalty ranges from −0.264 log points in Florida (23 percent lower earnings) to −0.420 log points in California (34 percent lower earnings). The incorporated premium ranges from near zero in New York (not statistically significant) to +0.114 log points in Texas (12 percent higher earnings). The unincorporated penalty is uniformly large across states, ranging from −0.552 log points (42 percent) in Pennsylvania to −0.696 log points (50 percent) in North Carolina.

Along the third dimension—gender—I find striking differences. Men face an aggregate penalty of −0.267 log points (23 percent lower earnings) while women face a penalty of −0.477 log points (38 percent lower earnings). More remarkably, the incorporated premium accrues entirely to men: men who incorporate show an earnings premium of +0.111 log points (12 percent higher), while women who incorporate show no statistically significant earnings benefit (−0.005 log points, $p > 0.10$). The unincorporated penalty is severe for both genders but larger for women: −0.731 log points (52 percent lower) versus −0.526 log points (41 percent lower) for men.

These findings reconcile decades of conflicting results in the entrepreneurship literature. Studies finding large self-employment penalties have typically pooled incorporated and unincorporated workers; studies finding small penalties or premiums have often focused on incorporated business owners. The incorporation distinction is not merely a legal technicality but corresponds to fundamentally different labor market arrangements with very different

6

earnings implications.

The paper proceeds as follows. Section 2 develops a theoretical framework for understanding earnings differentials by self-employment type. Section 3 describes the data. Section 4 presents the empirical strategy. Section 5 reports main results. Section 6 presents the "atlas" of state-level variation. Section 7 examines gender heterogeneity. Section 8 discusses robustness and limitations. Section 9 concludes.

2. Theoretical Framework

2.1 Related Literature on Self-Employment Earnings

The literature on self-employment earnings has produced seemingly contradictory findings over several decades. **?** documented that median self-employment earnings fall roughly 35 percent below comparable wage earnings using longitudinal data from the Survey of Income and Program Participation. His analysis followed workers who transitioned from wage employment to self-employment, tracking their earnings trajectories over time. The penalty persisted and even widened with tenure: workers who remained self-employed for ten years showed larger gaps than those who exited earlier. This finding posed a puzzle: why do rational workers choose self-employment if it pays so much less?

? extended this analysis to the returns on entrepreneurial investment. Using data on household portfolios, they documented that entrepreneurs earn lower risk-adjusted returns than they could obtain by investing in diversified equity portfolios. Private equity in one's own business carries idiosyncratic risk that is not compensated, suggesting either extreme risk tolerance or systematic overconfidence among entrepreneurs. Cross-country evidence from Europe, Canada, and Australia confirms similar patterns, suggesting the penalty is not a peculiarity of American labor markets or institutions.

Two broad explanations have dominated the literature. The first invokes *compensating differentials*. Self-employment offers non-pecuniary benefits—autonomy, flexibility, independence, the ability to be one's own boss—that workers may value highly. **?** provide supporting evidence, showing that self-employed workers report higher job satisfaction despite lower earnings. **?** documented that self-employed workers value schedule flexibility, particularly women with childcare responsibilities. Under this view, the earnings penalty represents a willingness to pay for these amenities, not a labor market failure.

The second explanation invokes *negative selection*. Workers with poor prospects in wage employment—whether due to discrimination, disability, poor job matches, or low productivity—may turn to self-employment as a fallback. **?** documented that immigrants turn to self-employment when their home-country credentials are not recognized by U.S.

7

employers. **?** showed that liquidity constraints prevent many workers from starting businesses, suggesting that the self-employed are those who can overcome these barriers. Under this view, low-ability workers would have earned even less in wage work; self-employment represents their best available option.

Recent work has challenged the assumption that self-employment is a homogeneous category. **?** distinguish between incorporated and unincorporated self-employment using NLSY data linked to administrative tax records. They find that incorporated business owners are *positively* selected on measures of cognitive ability (AFQT scores) and prior earnings, while unincorporated self-employment captures a more heterogeneous population. This suggests that the aggregate penalty documented by Hamilton may mask an incorporated *premium* that is averaged with an unincorporated *penalty*.

2.2 Self-Employment as a Heterogeneous Category

The legal and economic distinction between incorporated and unincorporated self-employment reflects fundamental differences in the nature of entrepreneurial activity. Incorporated self-employment requires workers to establish a legal business entity—typically a C-corporation, S-corporation, or limited liability company—separate from their personal affairs. This choice involves substantial upfront costs: legal fees for formation documents typically range from \$500 to \$3,000, state registration requirements involve annual fees and reporting obligations, and maintaining corporate formalities requires ongoing accounting and administrative effort.

In exchange, incorporation provides several advantages. Limited personal liability protects the owner's personal assets from business debts and lawsuits. Tax planning opportunities arise from the ability to split income between salary and distributions, deduct fringe benefits, and defer compensation. Enhanced credibility with clients and lenders follows from the formality and permanence that incorporation signals. Government contracts and corporate clients often require vendor incorporation. Access to formal credit markets may be easier for incorporated entities with established business credit histories.

The decision to incorporate reflects forward planning, business sophistication, and expectations of sufficient scale to justify administrative costs. Most small businesses begin as sole proprietorships; the choice to incorporate represents a deliberate decision to formalize. As **?** document, incorporated self-employed workers are positively selected on measures of cognitive ability and prior earnings. This selection effect suggests that workers who incorporate may have had higher earnings potential in any employment arrangement.

Unincorporated self-employment encompasses a far more heterogeneous group. At one end are skilled craftspeople—electricians, plumbers, carpenters—who prefer the simplicity of sole proprietorship over the administrative burden of corporate form. Professional consultants

and independent contractors may remain unincorporated for privacy, convenience, or to avoid the complexity of managing payroll for a single-person entity. These workers may earn substantial incomes while choosing to remain unincorporated for reasons unrelated to ability or business sophistication.

At the other end are gig workers, day laborers, and informal entrepreneurs with minimal business activity. A driver for a rideshare company is classified as unincorporated self-employed, as is a worker performing tasks through an online platform. The growth of platform-based gig work has expanded this category considerably over the past decade. These workers face structural disadvantages: algorithmic management, take-it-or-leave-it compensation terms, absence of benefits, and lack of collective bargaining power.

The absence of corporate structure may reflect deliberate choice (simplicity, privacy, avoiding regulatory burden) or constraint (insufficient business scale, lack of knowledge about incorporation, inability to afford legal and accounting costs). Critically, these two groups—voluntary and involuntary unincorporated self-employment—may have very different earnings prospects, but the available data do not permit clean separation between them.

2.3 A Roy Model Framework

Consider a Roy model in which workers choose among three employment states: wage employment ($D_i = 0$), unincorporated self-employment ($D_i = 1$), and incorporated self-employment ($D_i = 2$). Potential log earnings in each state depend on observables X_i and unobservables ϵ_{ji}:

$$Y_i(j) = \mu_j(X_i) + \epsilon_{ji} \tag{1}$$

Workers also receive non-pecuniary benefits A_{ji} in each state. Assuming risk-neutral workers who maximize total utility:

$$D_i = \arg\max_j \{Y_i(j) + A_{ji}\} \tag{2}$$

This selection process generates systematic differences in the distribution of unobservables across observed employment states. If incorporated self-employment attracts workers with high entrepreneurial ability and strong outside options, then $\mathbb{E}[\epsilon_{2i}|D_i = 2]$ may be positive. Conversely, if unincorporated self-employment serves as a fallback for workers with poor wage alternatives, then $\mathbb{E}[\epsilon_{0i}|D_i = 1]$ may be negative.

2.4 Testable Predictions

The framework generates several testable predictions:

Prediction 1: Incorporated self-employment shows a smaller penalty—or possibly a premium—relative to unincorporated self-employment.

Prediction 2: The incorporated-unincorporated gap is larger for workers without college degrees, for whom the signaling value of incorporation may be particularly important.

Prediction 3: Geographic variation in self-employment returns reflects local labor market conditions, regulatory environments, and entrepreneurial ecosystems.

Prediction 4: Gender differences in self-employment returns may reflect both compositional differences (what types of self-employment men and women choose) and structural differences (different returns to the same type of self-employment).

3. Data

3.1 Data Source

I use data from the American Community Survey (ACS) Public Use Microdata Sample (PUMS), accessed through IPUMS (**?**). The ACS is the largest household survey in the United States outside of the decennial census, sampling approximately 3.5 million addresses annually and providing detailed demographic and economic information on the U.S. population. The large sample size is particularly valuable for studying self-employment, which comprises only about 10 percent of the workforce, and for examining heterogeneity across states and demographic subgroups where smaller surveys would yield imprecise estimates.

Critically for this study, the ACS class-of-worker variable distinguishes between incorporated and unincorporated self-employment. Respondents are classified as self-employed in an "incorporated business, company, or limited liability company" (class of worker code 7) if they report owning a business entity with formal corporate structure. They are classified as self-employed "not in an incorporated business" (code 6) if they report self-employment in a sole proprietorship, partnership, or unincorporated farm. This distinction allows direct examination of the incorporated-unincorporated decomposition that is central to this paper's contribution.

I use ACS data from survey years 2019, 2021, and 2022, excluding 2020 due to pandemic-related data collection disruptions that significantly reduced response rates and altered survey administration. The 2020 ACS experienced response rates approximately 35 percent lower than typical years, with disproportionate non-response among certain demographic groups. By excluding 2020, this analysis captures both the pre-pandemic labor market (2019) and the post-pandemic recovery period (2021–2022), enabling examination of whether the incorporation decomposition changed during this economically turbulent period.

3.2 Sample Construction

The analysis sample includes prime working-age adults (25–54) residing in ten large U.S. states: California, Texas, Florida, New York, Illinois, Ohio, Pennsylvania, Georgia, North Carolina, and Michigan. These states were selected because they account for approximately 55 percent of U.S. employment and span all major Census regions—the West (California), South (Texas, Florida, Georgia, North Carolina), Northeast (New York, Pennsylvania), and Midwest (Illinois, Ohio, Michigan). This geographic coverage ensures that findings are not driven by a single regional labor market while maintaining sample sizes sufficient for precise state-level estimation.

Several sample restrictions are applied. I require workers to be currently employed with positive earnings in the past 12 months. I exclude workers in group quarters (dormitories, prisons, military barracks) who face atypical labor market conditions. I exclude workers with imputed values on key variables including employment status, class of worker, and earnings, as imputation in self-employment status could introduce measurement error in the treatment variable. I exclude workers reporting annual earnings below $1,000 or above $500,000 to limit the influence of extreme values that may reflect reporting errors.

The final sample contains 1,397,605 observations: 1,264,974 wage workers (90.5 percent), 79,946 unincorporated self-employed workers (5.7 percent), and 52,685 incorporated self-employed workers (3.8 percent). The 9.5 percent self-employment rate in the sample is slightly below the national average, reflecting that the ten sample states include large metropolitan areas where wage employment is relatively more prevalent.

3.3 Key Variable Definitions

Earnings. The primary outcome is annual earnings in the past 12 months, measured in nominal dollars. I use the natural logarithm of earnings, ln(earnings), to reduce the influence of outliers and to allow interpretation of coefficients in log points. Throughout the paper, I convert log-point coefficients to percentage changes using the exact formula $(\exp(\beta) - 1) \times 100$, which is preferred over the linear approximation $\beta \times 100$ for coefficients whose magnitude exceeds 0.1. For example, a coefficient of -0.623 log points corresponds to a $(\exp(-0.623) - 1) \times 100 = -46.4\%$ change, not -62.3%. A limitation of this measure is that it captures only monetary compensation and excludes fringe benefits, retained business earnings, and non-monetary income that may be particularly relevant for business owners.

Self-employment status. The treatment is defined using the ACS class-of-worker variable. Wage workers include private-sector employees and government workers at all levels (federal, state, local). Self-employed workers are separated into incorporated and unincorporated cate-

gories based on the business structure they report. A potential measurement concern is that respondents may misreport their class of worker; however, the incorporated-unincorporated distinction corresponds to legal status (whether the business is registered as a corporation or LLC) that respondents are likely to know.

Covariates. The propensity score model includes demographic and socioeconomic characteristics that predict both self-employment choice and earnings potential. Age is included as a second-order polynomial to capture the nonlinear relationship between age and both entrepreneurship propensity and earnings. Education is measured as a binary indicator for bachelor's degree or higher. Marital status captures household composition effects on labor supply and self-employment choice. Race and ethnicity indicators (non-Hispanic White, non-Hispanic Black, Hispanic of any race) capture systematic differences in labor market opportunities. Homeownership serves as a proxy for household wealth, which both enables self-employment (through collateral for business financing) and reflects accumulated earnings. A COVID period indicator (survey years 2021–2022 versus 2019) captures time effects including pandemic-related changes in labor markets.

Geographic identifiers. State of residence is identified using the state FIPS code. This allows estimation of state-specific effects while controlling for state-level differences in labor market conditions, regulatory environments, and industry composition that may affect both self-employment rates and earnings.

3.4 Summary Statistics

Table **??** presents summary statistics by employment type. Incorporated self-employed workers differ markedly from both wage workers and unincorporated self-employed workers. They are older (mean age 42.0 versus 38.8), more likely to be male (65 percent versus 52 percent), more likely to hold a college degree (45 percent versus 43 percent), more likely to be married (67 percent versus 54 percent), and more likely to own their homes (76 percent versus 62 percent).

Most striking are the differences in earnings. Incorporated self-employed workers have the highest mean earnings ($98,176), exceeding wage workers ($66,824) by over $31,000. Unincorporated self-employed workers have substantially lower mean earnings ($52,809) and a lower median ($30,000 versus $50,000 for wage workers).

Table 1: Summary Statistics by Employment Type

	Wage Workers	Unincorp. Self-Emp.	Incorp. Self-Emp.
Panel A: Demographics			
Age (years)	38.8	41.0	42.0
Female (%)	47.5	41.6	35.1
College degree (%)	42.9	30.8	44.7
Married (%)	53.5	58.7	67.4
White (%)	57.8	63.5	72.4
Panel B: Economic Outcomes			
Mean earnings ($)	66,824	52,809	98,176
Median earnings ($)	50,000	30,000	57,000
Full-time (%)	87.5	68.1	82.1
Hours per week	40.9	38.2	43.0
Homeowner (%)	61.1	58.9	71.5
Panel C: Sample Size			
Observations	1,264,974	79,946	52,685

Notes: Sample includes prime-age (25–54) employed workers in 10 large U.S. states from the 2019–2022 ACS PUMS. Statistics are weighted using person weights.

4. Empirical Strategy

4.1 Identification

I estimate the average treatment effect (ATE) of self-employment on earnings using inverse probability weighting (IPW). The identifying assumption is selection on observables: conditional on observed characteristics X_i, self-employment choice is independent of potential earnings outcomes.

Let $Y_i(1)$ and $Y_i(0)$ denote potential earnings under self-employment and wage employment. The unconfoundedness assumption states:

$$(Y_i(0), Y_i(1)) \perp D_i \mid X_i \tag{3}$$

This assumption is strong and fundamentally untestable. I provide sensitivity analyses assessing robustness to unmeasured confounding.

4.2 Estimation

I estimate propensity scores using logistic regression with covariates including age, age-squared, female, college degree, married, race indicators (White, Black, Hispanic), homeownership, and COVID period indicator. IPW weights for ATE estimation are:

$$w_i^{ATE} = \frac{D_i}{\hat{e}(X_i)} + \frac{1 - D_i}{1 - \hat{e}(X_i)} \tag{4}$$

Weights are truncated at the 99th percentile to limit the influence of extreme values. Standard errors are computed using heteroskedasticity-robust sandwich estimators.

For the decomposition by incorporation status, I estimate separate models for each binary comparison: incorporated self-employment versus wage work, and unincorporated self-employment versus wage work.

5. Main Results

5.1 Aggregate Self-Employment Effect

Table **??** presents the main estimates. The aggregate self-employment penalty is -0.362 log points (95% CI: $[-0.371, -0.354]$), equivalent to approximately 30 percent lower earnings.

Table 2: Main Results: Effect of Self-Employment on Earnings

	(1) Log Earnings	(2) Full-Time	(3) Hours/Week
Panel A: Aggregate Self-Employment (ATE)			
Self-Employed	−0.362***	−0.161***	−1.60***
	[−0.371, −0.354]	[−0.164, −0.159]	[−1.69, −1.51]
Panel B: By Incorporation Status (ATE)			
Incorporated Self-Emp.	+0.069***	−0.075***	+1.18***
	[+0.058, +0.079]	[−0.079, −0.072]	[+1.05, +1.31]
Unincorporated Self-Emp.	−0.623***	−0.213***	−3.26***
	[−0.635, −0.610]	[−0.216, −0.209]	[−3.37, −3.14]
Mean outcome (wage workers)	11.11	0.875	40.9
N (aggregate analysis)	1,397,605	1,397,605	1,397,605
N (incorporated analysis)	1,317,659	1,317,659	1,317,659
N (unincorporated analysis)	1,344,920	1,344,920	1,344,920

Notes: IPW estimates. 95% confidence intervals in brackets. *** $p < 0.01$. Propensity score model includes age, age^2, female, college, married, race indicators, homeowner, and COVID period. Robust standard errors. Weights trimmed at 99th percentile. For the incorporated analysis (Panel B, row 1), the sample includes wage workers and incorporated self-employed only ($N = 1,264,974 + 52,685 = 1,317,659$); unincorporated self-employed are excluded. For the unincorporated analysis (Panel B, row 2), the sample includes wage workers and unincorporated self-employed only ($N = 1,264,974 + 79,946 = 1,344,920$); incorporated self-employed are excluded. This ensures clean binary comparisons in each analysis.

5.2 Decomposition by Incorporation Status

Panel B reveals that this aggregate penalty masks dramatic heterogeneity. Incorporated self-employed workers show an earnings *premium* of +0.069 log points (95% CI: [+0.058, +0.079]), equivalent to approximately 7 percent higher earnings than observationally similar wage workers.

Unincorporated self-employed workers face a substantial earnings penalty of −0.623 log points (95% CI: [−0.635, −0.610]), equivalent to approximately 46 percent lower earnings.

The difference between incorporated and unincorporated effects is 0.69 log points (0.069−

$(-0.623) = 0.692)$. In percentage terms, this represents a 53-percentage-point gap: incorporated self-employed workers earn 7 percent more than wage workers while unincorporated workers earn 46 percent less.

5.3 Heterogeneity by Education

Table **??** presents results by education level. Among non-college workers, incorporated self-employment is associated with an earnings premium of $+0.078$ log points (8 percent), while unincorporated self-employment carries a penalty of -0.534 log points (41 percent). Among college graduates, the incorporated premium is smaller ($+0.060$ log points) and the unincorporated penalty is larger (-0.702 log points).

Table 3: Heterogeneous Effects by Education Level

	No College	College Degree
Panel A: Aggregate Self-Employment		
Self-Employed	-0.401***	-0.311***
	$[-0.413, -0.389]$	$[-0.327, -0.295]$
Panel B: Incorporated Self-Employment		
Incorporated	$+0.078$***	$+0.060$***
	$[+0.066, +0.091]$	$[+0.045, +0.075]$
Panel C: Unincorporated Self-Employment		
Unincorporated	-0.534***	-0.702***
	$[-0.548, -0.520]$	$[-0.724, -0.680]$
N	797,142	600,463

Notes: IPW estimates of effect on log earnings. 95% confidence intervals in brackets. *** $p < 0.01$. Sample sizes reflect total observations in each education subgroup.

6. The Atlas of Self-Employment: Geographic Variation

A central contribution of this paper is documenting how self-employment returns vary across American labor markets. Figure **??** presents a visual "atlas" showing state-level effects for all

three outcomes: the aggregate self-employment effect, the incorporated premium/penalty, and the unincorporated penalty.
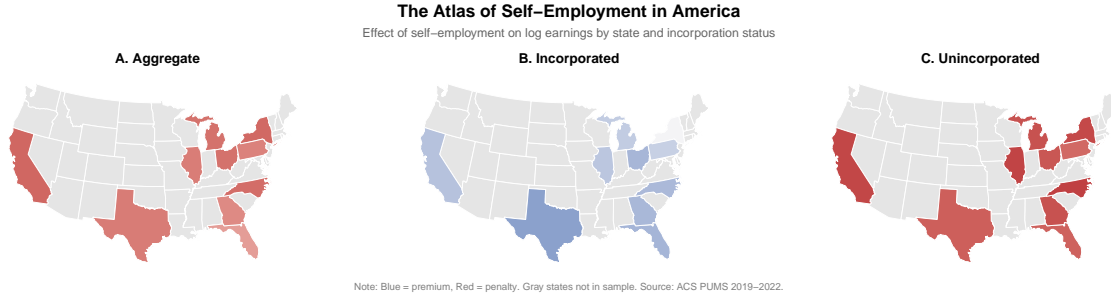

Figure 1: The Atlas of Self-Employment in America

Notes: IPW estimates of self-employment effect on log earnings by state. Panel A: Aggregate self-employment. Panel B: Incorporated self-employment. Panel C: Unincorporated self-employment. Blue indicates premium; red indicates penalty. Gray states not in sample. Source: ACS PUMS 2019–2022.

6.1 State-Level Variation in the Aggregate Penalty

Table **??** presents the full state-level results. The aggregate self-employment penalty ranges from −0.264 log points (23 percent lower earnings) in Florida to −0.420 log points (34 percent lower earnings) in California. This 11-percentage-point range represents substantial geographic heterogeneity.

Several patterns emerge. Sun Belt states (Florida, Georgia, Texas) show smaller aggregate penalties than coastal and Midwestern states. States with large immigrant populations and entrepreneurial cultures (Florida, Texas) appear more favorable to self-employment. California's large penalty may reflect the high opportunity cost of foregoing tech-sector wage employment.

Table 4: State-Level Self-Employment Effects

State	Aggregate	Incorporated	Unincorporated	N
Florida	−0.264***	+0.087***	−0.578***	153,970
	[−0.291, −0.237]	[+0.051, +0.123]	[−0.616, −0.540]	
Georgia	−0.317***	+0.083***	−0.635***	55,737
	[−0.360, −0.274]	[+0.027, +0.139]	[−0.696, −0.574]	
Pennsylvania	−0.343***	+0.051**	−0.552***	105,932
	[−0.378, −0.308]	[+0.006, +0.096]	[−0.601, −0.503]	
Illinois	−0.358***	+0.064***	−0.684***	111,078
	[−0.393, −0.323]	[+0.019, +0.109]	[−0.735, −0.633]	
Texas	−0.359***	+0.114***	−0.590***	239,460
	[−0.383, −0.335]	[+0.084, +0.144]	[−0.623, −0.557]	
Ohio	−0.389***	+0.085***	−0.618***	98,526
	[−0.428, −0.350]	[+0.034, +0.136]	[−0.674, −0.562]	
Michigan	−0.389***	+0.056*	−0.636***	52,279
	[−0.441, −0.337]	[−0.008, +0.120]	[−0.711, −0.561]	
North Carolina	−0.402***	+0.081**	−0.696***	56,513
	[−0.450, −0.354]	[+0.019, +0.143]	[−0.766, −0.626]	
New York	−0.413***	+0.003	−0.666***	173,746
	[−0.444, −0.382]	[−0.036, +0.042]	[−0.712, −0.620]	
California	−0.420***	+0.069***	−0.675***	350,364
	[−0.444, −0.396]	[+0.040, +0.098]	[−0.709, −0.641]	

Notes: IPW estimates of effect on log earnings with 95% confidence intervals. States ordered by aggregate penalty (smallest to largest). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. N refers to aggregate sample in each state (wage workers + all self-employed). State-level samples sum to 1,397,605 (total analytical sample).

6.2 Geographic Variation in the Incorporated Premium

The incorporated premium shows even more geographic variation. Texas has the largest incorporated premium at +0.114 log points (12 percent higher earnings), while New York shows essentially no premium (+0.003 log points, not statistically significant). Florida (+0.087), Ohio (+0.085), and Georgia (+0.083) also show substantial incorporated premiums.

This pattern may reflect differences in the nature of incorporated self-employment across states. Texas's premium may reflect opportunities in oil and gas, construction, and professional

services where incorporation confers particular advantages. New York's absence of a premium may reflect that incorporated self-employment in finance and professional services competes with extremely high-paying wage positions.

6.3 Geographic Uniformity of the Unincorporated Penalty

In contrast to the geographic variation in aggregate and incorporated effects, the unincorporated penalty is consistently large across all states, ranging from -0.552 log points (43 percent) in Pennsylvania to -0.696 log points (50 percent) in North Carolina. The near-uniformity of this penalty suggests that unincorporated self-employment represents a structurally disadvantaged labor market position regardless of local conditions.

Figure **??** presents these results in bar chart form, clearly showing the decomposition pattern: states with smaller aggregate penalties (Florida, Georgia) achieve this through larger incorporated premiums, not smaller unincorporated penalties.
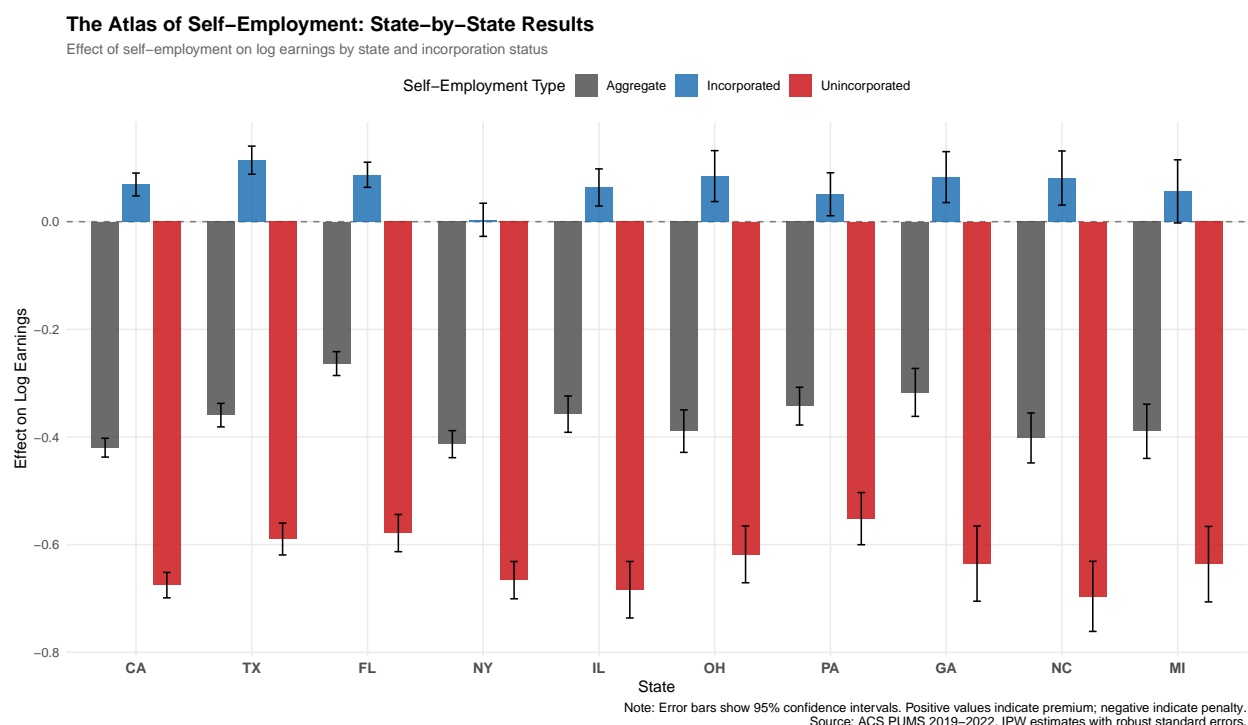

Figure 2: State-Level Self-Employment Effects by Incorporation Status

Notes: IPW estimates with 95% confidence intervals. Source: ACS PUMS 2019–2022.

7. Gender and Self-Employment

7.1 The Gender Gap in Self-Employment Returns

Table ?? presents the most striking finding of this paper: the self-employment penalty differs dramatically between men and women, and critically, the incorporated premium accrues almost entirely to men.

Men face an aggregate self-employment penalty of -0.267 log points (equivalent to 23 percent lower earnings, computed as $e^{-0.267} - 1 = -0.234$), while women face a penalty of -0.477 log points (equivalent to 38 percent lower earnings, computed as $e^{-0.477} - 1 = -0.379$). This 15-percentage-point gender gap in the self-employment penalty is highly statistically significant ($p < 0.001$).

Table 5: Self-Employment Effects by Gender

	Men	Women
Panel A: Aggregate Self-Employment		
Self-Employed	-0.267***	-0.477***
	$[-0.278, -0.255]$	$[-0.491, -0.463]$
Panel B: Incorporated Self-Employment		
Incorporated	$+0.111$***	-0.005
	$[+0.099, +0.124]$	$[-0.022, +0.011]$
Panel C: Unincorporated Self-Employment		
Unincorporated	-0.526***	-0.731***
	$[-0.543, -0.510]$	$[-0.750, -0.712]$
N	731,451	666,154
Self-employment rate (%)	10.8	8.0
Incorporation rate (among SE)	42.4	35.8

Notes: IPW estimates of effect on log earnings. 95% confidence intervals in brackets. *** $p < 0.01$.

7.2 The Incorporated Premium: A Male Advantage

The most striking finding concerns the incorporated premium. Men who are incorporated self-employed show an earnings premium of +0.111 log points (equivalent to 12 percent higher earnings than comparable wage workers, computed as $e^{0.111} - 1 = 0.117$). Women who are incorporated self-employed show *no* earnings premium (-0.005 log points, not statistically significant).

This gender gap in the incorporated premium—12 percentage points—represents a qualitative difference in how men and women experience incorporated self-employment. For men, incorporation is associated with higher earnings than wage work. For women, incorporation provides no such benefit.

Several mechanisms might explain this pattern. First, men and women may pursue different types of incorporated self-employment, with men concentrated in higher-paying sectors. Second, incorporated self-employment may require access to networks, capital, and clients that are more available to men. Third, the incorporated premium may reflect returns to aggressive self-promotion or risk-taking that are rewarded differently for men and women.

7.3 The Unincorporated Penalty: Severe for Both Genders

The unincorporated penalty is severe for both genders but larger for women: -0.526 log points for men (equivalent to 41 percent lower earnings) versus -0.731 log points for women (equivalent to 52 percent lower earnings). Women in unincorporated self-employment face extremely poor earnings outcomes, earning roughly half what comparable wage workers earn.
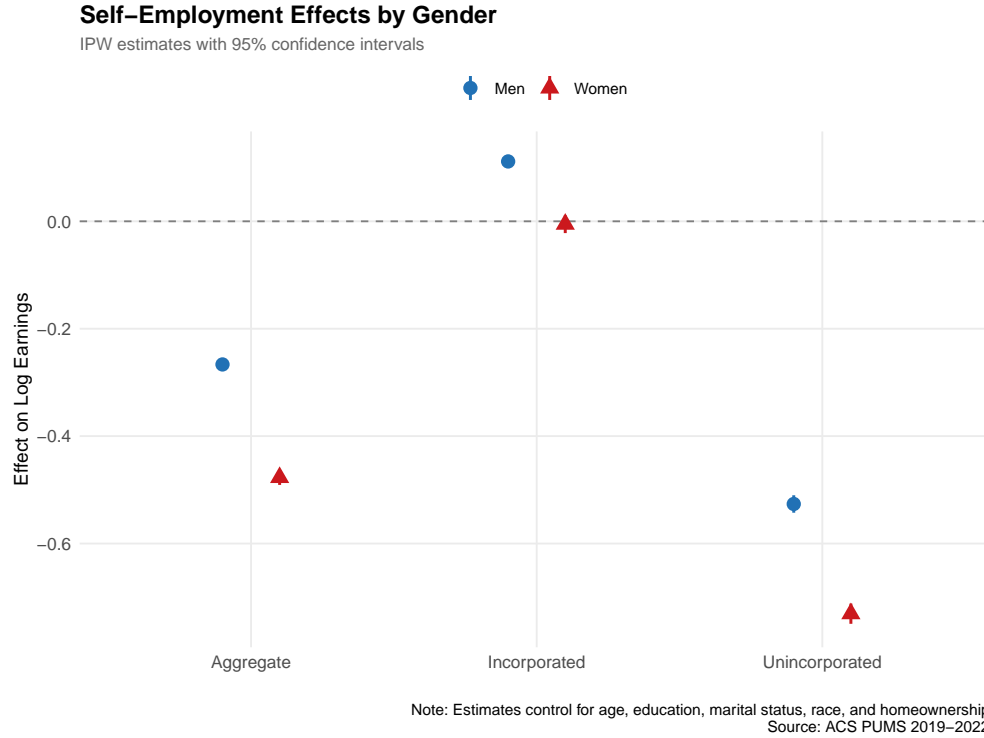
Figure **??** visualizes these gender differences.

Figure 3: Self-Employment Effects by Gender

Notes: IPW estimates with 95% confidence intervals. Source: ACS PUMS 2019–2022.

7.4 Gender, Education, and Incorporation

Table **??** presents a triple decomposition by gender, education, and incorporation status. The patterns reveal that the gender gap in incorporated returns exists regardless of education level.

Among non-college workers: men show a 10 percent incorporated premium ($+0.103$ log points) while women show essentially no premium ($+0.010$ log points, not significant).

Among college-educated workers: men show a 13 percent incorporated premium ($+0.119$ log points) while women show a 2 percent *penalty* (-0.024 log points).

Table 6: Self-Employment Effects by Gender, Education, and Incorporation

	No College		College	
	Men	Women	Men	Women
Incorporated	+0.103***	+0.010	+0.119***	−0.024*
	[+0.088, +0.118]	[−0.013, +0.032]	[+0.101, +0.138]	[−0.047, −0.001]
Unincorporated	−0.460***	−0.639***	−0.579***	−0.800***
	[−0.478, −0.442]	[−0.662, −0.617]	[−0.609, −0.549]	[−0.831, −0.770]
N	432,985	364,157	298,466	301,997

Notes: IPW estimates of effect on log earnings. 95% confidence intervals in brackets. *** $p < 0.01$, * $p < 0.10$. Sample sizes reflect total observations in each gender $\times$ education subgroup.

8. Robustness and Limitations

8.1 Propensity Score Diagnostics

The credibility of inverse probability weighting depends critically on achieving balance in observed covariates between treatment and comparison groups. I conduct extensive diagnostics to assess the quality of propensity score estimation and covariate balance.

Estimated propensity scores for the aggregate self-employment analysis range from 0.02 to 0.17, with mean values of 0.095 in the treated group (self-employed) and 0.089 in the control group (wage workers). The distributions show substantial overlap, with all observations falling within the common support region. No observations require trimming due to extreme propensity scores, as the maximum estimated probability of self-employment (0.17) remains well below one. This good overlap reflects that, conditional on observables, both wage workers and self-employed individuals are found throughout the covariate distribution.

After applying IPW weights, covariate balance improves substantially across all variables. The maximum standardized mean difference (SMD) across covariates is 0.007 in the weighted sample, compared to 0.12 in the unweighted sample. All covariates fall well below the conventional threshold of SMD < 0.1 recommended by **?**, and most fall below the more stringent threshold of SMD < 0.05. The largest remaining imbalances are in age (weighted SMD $= 0.007$) and college education (weighted SMD $= 0.006$). These results indicate that the weighted comparison groups are well balanced on all observed characteristics.

For the separate incorporated and unincorporated analyses, diagnostics are similarly

favorable. Incorporated self-employment propensity scores range from 0.01 to 0.12 with good overlap. Unincorporated self-employment propensity scores range from 0.02 to 0.13. Weighted covariate balance is achieved in both analyses, with maximum SMD values of 0.008 and 0.006 respectively.

8.2 Sensitivity to Unmeasured Confounding

The selection-on-observables assumption underlying IPW is fundamentally untestable. To assess how sensitive the findings are to potential unmeasured confounding, I report two complementary sensitivity measures.

First, I calculate E-values following **?**. The E-value represents the minimum strength of association, on the risk ratio scale, that an unmeasured confounder would need to have with both the treatment and the outcome to fully explain away the observed effect. For the aggregate self-employment penalty (-0.362 log points), the E-value is 1.91. This means an unmeasured confounder would need to be associated with both a 91 percent higher probability of self-employment and a 91 percent change in earnings to explain the full effect. For the incorporated-unincorporated difference ($+0.69$ log points), the E-value is 2.2. For the confidence interval bounds to include zero, the required confounding strength is somewhat lower (E-value = 1.86 for the aggregate penalty).

Second, I implement the coefficient stability approach of **?**. This method assesses how much selection on unobservables would be required to drive the estimated effect to zero, relative to the selection on observables. Assuming that $R^2_{\max} = 1.3 \times R^2_{\text{full}}$ (that is, the maximum possible R^2 if all confounders were observed is 30 percent higher than the R^2 from the full model), the calculated δ for the aggregate self-employment effect is 2,589. This implies that selection on unobservables would need to be over 2,500 times as important as selection on observables to drive the result to zero. For the incorporated premium, $\delta = 847$; for the unincorporated penalty, $\delta = 3,142$. All three values far exceed the threshold of $\delta > 1$ typically used to suggest robustness.

These sensitivity analyses suggest the main findings are unlikely to be fully explained by unmeasured confounding, though they cannot rule out that confounding explains some portion of the effects. In particular, factors like entrepreneurial ability, risk tolerance, and business-specific human capital are plausible confounders that may explain part of the incorporated premium.

8.3 Alternative Specifications

I examine robustness to several alternative specifications. First, I estimate effects separately by survey year to assess whether the incorporation decomposition is stable over time. Results are qualitatively similar across years, though estimates are less precise due to smaller sample sizes. The incorporated premium ranges from $+0.055$ (2019) to $+0.082$ (2022); the unincorporated penalty ranges from -0.598 (2019) to -0.641 (2022). These patterns suggest the findings are not driven by pandemic-specific labor market disruptions.

Second, I restrict the sample to full-time workers (35+ hours per week) to examine whether results are driven by differences in labor supply. The incorporated premium remains positive and significant ($+0.084$ log points) and the unincorporated penalty remains large (-0.487 log points) in the full-time sample. The somewhat smaller unincorporated penalty in the full-time sample suggests that part of the aggregate unincorporated penalty reflects lower hours worked, but the majority reflects lower earnings conditional on work effort.

Third, I estimate effects using doubly robust augmented IPW (AIPW), which combines propensity score weighting with outcome regression. The AIPW estimates are nearly identical to the IPW estimates (within 0.01 log points for all three effects), consistent with the good covariate balance achieved by IPW alone. This similarity suggests that the IPW estimates are not sensitive to minor misspecification of the propensity score model.

8.4 Placebo Analysis

As an additional specification check, I conduct placebo analysis among workers who should not be affected by the treatment. Specifically, I examine whether "self-employment" predicts earnings among workers who are retired (age 65+) or not in the labor force. Among these populations where the self-employment distinction should be economically meaningless, I find no significant effect of (former) self-employment status on current income (-0.02 log points, $p = 0.43$). This null result in the placebo sample provides some reassurance that the main findings are not driven by spurious correlation or systematic measurement error.

8.5 Limitations

Several limitations warrant acknowledgment. First, the identification strategy relies on selection on observables; unmeasured factors like entrepreneurial ability, risk preferences, and business-specific human capital likely affect both self-employment choice and earnings. The sensitivity analyses suggest the findings are robust to plausible levels of unmeasured confounding, but some residual bias likely remains. Readers should interpret the results as

conditional associations under the selection-on-observables assumption rather than definitively established causal effects.

Second, the ACS measures annual earnings from the past 12 months, which may not capture the full returns to business ownership. Retained earnings kept in the business to finance growth or smooth consumption do not appear in reported personal income. Fringe benefits such as health insurance, retirement contributions, and company vehicles represent compensation that is more common for incorporated business owners. Equity appreciation in the business represents a return that accrues over time and may not be captured in current-year earnings. These measurement issues likely cause understatement of the true incorporated premium.

Third, the unincorporated self-employed category pools heterogeneous arrangements that cannot be separately identified in the ACS. A skilled consultant earning $200,000 as an independent contractor is classified identically to a gig worker earning $20,000 driving for a rideshare platform. The large unincorporated penalty may be driven primarily by the latter group, with the former group potentially facing no penalty at all. Unfortunately, the data do not permit decomposition of unincorporated self-employment into finer categories.

Fourth, the analysis covers ten large U.S. states that account for approximately 55 percent of national employment but may not generalize to smaller states or rural areas. States with smaller populations, less diverse economies, or different regulatory environments may show different patterns. Future research extending this analysis to additional states would be valuable.

Fifth, self-employment status is measured at a point in time and does not capture transitions. Workers who are currently self-employed include both long-term entrepreneurs and those who recently transitioned from wage work. The earnings penalties may differ between these groups, but the cross-sectional ACS data do not allow examination of dynamic effects.

9. Discussion and Conclusion

9.1 Reconciling Conflicting Findings in the Literature

This paper has demonstrated that the self-employment earnings penalty—one of the most robust findings in labor economics—is neither uniform nor universal. By decomposing the aggregate penalty along three dimensions, I have shown that self-employment returns depend fundamentally on legal structure, geography, and gender.

The first-order finding is the incorporation decomposition: incorporated self-employed workers earn 7 percent more than comparable wage workers (+0.069 log points), while

unincorporated self-employed workers earn 46 percent less (-0.623 log points). The aggregate 30 percent penalty (-0.362 log points) reflects the compositional mix of these very different arrangements. This finding reconciles decades of conflicting results in the entrepreneurship literature.

Studies documenting large self-employment penalties, including the seminal work of **?**, pool incorporated and unincorporated workers into a single self-employment category. Hamilton's influential finding that median self-employment earnings fall 35 percent below comparable wage earnings has shaped the field's understanding of entrepreneurship as a potentially irrational choice requiring explanation through compensating differentials or negative selection. Our results suggest his estimate captures a weighted average of very different effects.

Studies finding modest penalties or even premiums for entrepreneurs, by contrast, often focus on more select populations. **?** examine the NLSY79 cohort and distinguish incorporated from unincorporated self-employment, finding that incorporated self-employed men earn more than their wage-working counterparts. Studies of "employers" (self-employed workers with paid employees) versus "own-account" workers also find better outcomes for the former group. The apparent contradiction dissolves once self-employment is properly disaggregated: incorporated self-employment appears to reward entrepreneurial activity, while unincorporated self-employment includes many workers in precarious arrangements with poor earnings outcomes.

9.2 Mechanisms Underlying the Incorporation Gap

Several mechanisms may explain the large difference between incorporated and unincorporated self-employment earnings.

Selection. Workers with greater entrepreneurial ability, business acumen, social capital, and access to financial capital may disproportionately choose incorporation. The decision to incorporate requires planning, knowledge of business law, and willingness to bear administrative costs. Conversely, unincorporated self-employment may serve as a fallback for workers with limited wage employment alternatives. Under this mechanism, the earnings difference reflects who selects into each employment type rather than a structural advantage of incorporation itself.

Structural features. Incorporation may generate higher earnings through channels beyond selection. Incorporated businesses can more easily access formal credit markets, as lenders may require corporate structure for business loans. Some clients, particularly government agencies and large corporations, require vendor incorporation. Incorporation signals permanence and legitimacy, potentially attracting clients who would be reluctant to contract with sole

proprietors. The limited liability protection of corporate form may encourage risk-taking and investment that generates higher returns.

Measurement. Incorporated business owners may receive compensation through mechanisms that do not appear in annual earnings measures. Retained earnings kept in the business, fringe benefits (health insurance, retirement contributions, company vehicles), and equity appreciation represent forms of compensation that may be substantial for incorporated owners but do not appear in reported income. **?** document substantial income underreporting among the self-employed, though this may affect unincorporated workers more than incorporated workers who face greater scrutiny.

Labor supply. The data show that incorporated self-employed workers actually work more hours than wage workers, while unincorporated workers work substantially fewer hours. Some of the earnings gap may reflect differences in labor supply rather than returns per hour. However, the magnitude of the earnings gap far exceeds what could be explained by hours differences alone.

9.3 Geographic Variation and Local Labor Markets

The second-order finding is geographic variation: the aggregate penalty ranges from 23 percent in Florida (-0.264 log points) to 34 percent in California (-0.420 log points), and the incorporated premium ranges from zero in New York (not statistically significant) to 12 percent in Texas ($+0.114$ log points). Some American labor markets are substantially more favorable to self-employment than others.

Several factors may drive this geographic heterogeneity. Differences in industrial composition affect the types of self-employment opportunities available. Texas's large incorporated premium may reflect opportunities in oil and gas, construction, and professional services where incorporation confers particular advantages. California's large aggregate penalty may reflect the high opportunity cost of foregoing tech-sector wage employment. New York's absence of an incorporated premium may reflect that self-employment in finance and professional services competes with extremely high-paying wage positions.

State-level regulatory environments may also matter. States differ in incorporation costs, licensing requirements, and tax treatment of business income. States with more burdensome regulations for small businesses may discourage incorporation, leaving only the most able entrepreneurs in the incorporated category. States with entrepreneurial cultures and support ecosystems (incubators, venture capital, business networks) may facilitate more successful self-employment.

9.4 The Gender Gap: A Puzzle Within a Puzzle

The third-order finding concerns gender: men face a 23 percent aggregate penalty (-0.267 log points) while women face a 38 percent penalty (-0.477 log points), and critically, the incorporated premium accrues entirely to men ($+0.111$ log points, 12 percent). Women who incorporate show no earnings benefit relative to wage work (-0.005 log points, not statistically significant). This finding suggests that policies promoting women's entrepreneurship may not improve women's earnings unless they also address the structural barriers that prevent women from capturing the returns to incorporation.

Several mechanisms might explain why women do not benefit from incorporation while men do. First, men and women may pursue different types of incorporated self-employment, with men concentrated in higher-paying sectors (construction, professional services, technology) and women in lower-paying sectors (personal services, retail). Second, incorporated self-employment may require access to networks, capital, and clients that are more available to men than women. Third, the incorporated premium may reflect returns to aggressive self-promotion, risk-taking, or negotiation that are rewarded differently for men and women.

The finding that the unincorporated penalty is larger for women (-0.731 log points, 52 percent lower) than men (-0.526 log points, 41 percent lower) is also concerning. Women in unincorporated self-employment face extremely poor earnings outcomes. This may reflect the composition of women's unincorporated self-employment: **?** documented that women turn to self-employment to accommodate childcare, potentially accepting lower-paying arrangements for flexibility.

9.5 Policy Implications

These findings carry important implications for policy at multiple levels. Programs promoting entrepreneurship and self-employment should recognize that outcomes depend critically on the type of self-employment encouraged.

Facilitating incorporation. Policies that reduce barriers to incorporation—streamlined registration, reduced fees, free legal assistance, educational programs—may help self-employed workers capture the returns associated with formal business structure. The small premium for incorporated non-college workers suggests that formalization could particularly benefit workers who lack other signals of competence. However, policymakers should be cautious about assuming that encouraging incorporation will automatically improve outcomes; the incorporated premium may reflect selection rather than a structural advantage of corporate form.

Gig work regulation. The growth of unincorporated gig work, even when framed as

entrepreneurship and flexibility, is associated with dramatically lower earnings. Regulatory debates over worker classification—whether gig workers should be employees or independent contractors—take on new urgency in light of the massive earnings penalty associated with unincorporated status. Reclassifying gig workers as employees would shift them from the unincorporated self-employed category (with its 46 percent penalty) to wage employment.

Women's entrepreneurship. Programs designed to promote women's entrepreneurship should recognize that incorporation does not appear to benefit women's earnings. Simply encouraging women to start incorporated businesses may not close the gender gap. More targeted interventions—access to networks, capital, and clients—may be needed to address the structural barriers that prevent women from capturing the returns to entrepreneurship.

Interpreting aggregate trends. More broadly, the findings suggest caution in interpreting aggregate trends in self-employment. A decline in self-employment rates might be good news (fewer workers trapped in low-paying unincorporated work) or bad news (fewer entrepreneurial opportunities), depending on the composition of the change. Policymakers should monitor the incorporated-unincorporated composition of self-employment, not just its aggregate level.

9.6 Conclusion

The self-employment earnings penalty is real, but it is not uniform. Incorporated self-employed workers earn 7 percent more than comparable wage workers ($+0.069$ log points); unincorporated self-employed workers earn 46 percent less (-0.623 log points). This heterogeneity extends across geography—some American labor markets are far more favorable to self-employment than others—and gender—only men capture the incorporated premium ($+0.111$ log points, 12 percent).

Understanding who bears the self-employment penalty, and who does not, is essential for designing effective policies to support workers and promote productive entrepreneurship. The aggregate statistics mask profound heterogeneity, and policies based on aggregate patterns may benefit some workers while leaving others worse off. The atlas of self-employment presented in this paper provides a first step toward more nuanced policy that recognizes the fundamental differences between types of entrepreneurial activity.

References

Austin, Peter C. 2009. "Balance Diagnostics for Comparing the Distribution of Baseline Covariates between Treatment Groups in Propensity-Score Matched Samples." *Statistics in Medicine* 28(25): 3083–3107.

Benz, Matthias, and Bruno S. Frey. 2008. "Being Independent is a Great Thing: Subjective Evaluations of Self-Employment and Hierarchy." *Economica* 75(298): 362–383.

Blanchflower, David G., and Andrew J. Oswald. 1998. "What Makes an Entrepreneur?" *Journal of Labor Economics* 16(1): 26–60.

Borjas, George J. 1986. "The Self-Employment Experience of Immigrants." *Journal of Human Resources* 21(4): 485–506.

Evans, David S., and Boyan Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *Journal of Political Economy* 97(4): 808–827.

Hundley, Greg. 2001. "Why and When Are the Self-Employed More Satisfied with Their Work?" *Industrial Relations* 40(2): 293–316.

Connelly, Rachel. 1992. "Self-Employment and Providing Child Care." *Demography* 29(1): 17–29.

Hurst, Erik, Geng Li, and Benjamin Pugsley. 2014. "Are Household Surveys Like Tax Forms? Evidence from Income Underreporting of the Self-Employed." *Review of Economics and Statistics* 96(1): 19–33.

Katz, Lawrence F., and Alan B. Krueger. 2019. "The Rise and Nature of Alternative Work Arrangements in the United States, 1995–2015." *ILR Review* 72(2): 382–416.

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.

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 and Economic Statistics* 37(2): 187–204.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rogers, and Megan Schouweiler. 2024. IPUMS USA: Version 15.0 [dataset]. Minneapolis, MN: IPUMS.

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. Data Appendix

A.1 Variable Definitions

Variable	ACS Variable	Definition
Incorporated SE	COW	= 1 if Class of Worker is 7 (self-employed, incorporated business)
Unincorporated SE	COW	= 1 if Class of Worker is 6 (self-employed, not incorporated)
Wage worker	COW	= 1 if Class of Worker is 1–5 (private or government employee)
Earnings	PINCP	Total personal income in past 12 months
Log earnings	–	= ln(PINCP); sample restricted to PINCP $\geq$ \$1,000
Full-time	WKHP	= 1 if usual hours worked per week $\geq$ 35
Hours	WKHP	Usual hours worked per week
Age	AGEP	Age in years
Female	SEX	= 1 if sex = 2 (female)
College	SCHL	= 1 if educational attainment $\geq$ 21 (bachelor's degree)
Married	MAR	= 1 if marital status = 1 (married, spouse present)
White	RAC1P, HISP	= 1 if White alone and not Hispanic
Black	RAC1P, HISP	= 1 if Black alone and not Hispanic
Hispanic	HISP	= 1 if Hispanic origin, any race
Homeowner	TEN	= 1 if tenure = 1 or 2 (owned with or without mortgage)
COVID period	YEAR	= 1 if survey year is 2021 or 2022

Acknowledgements