The Effect of Industry on Individual Income: Evidence from U.S. Microdata

Author: Jonah Kress
Institution: Purdue University

---

1. Introduction

Industry choice plays a central role in determining individual income, reflecting both the structure of labor demand and the skills valued in different sectors. Understanding how industries differ in wage outcomes can offer insights on economic inequality and its influencing factors.

This analysis examines how total personal income varies across industries in the United States, with particular focus on the technical consulting sector. Technical consulting has become an increasingly prominent career path for graduates in economics, finance, and related fields. The sector's emphasis on "hard-coded" analytical skills and problem-solving may lead to higher compensation than other industries, even after controlling for confounding variables.

The primary objective of this report is to assess the potential salary difference between the technical consulting industry and the average of all other IPUMS-listed industries using a regression framework.

---

2. Data and Variables

The data used in this analysis come from the **Integrated Public Use Microdata Series (IPUMS)**, which provides anonymized microdata from the U.S. Census and the American Community Survey. The sample consists of adults (ages 18–64) who reported information for the year 2022.

Dependent Variable

- *Total Annual Personal Income (*coded as *'inctot'):* Annual income from all sources, expressed in 2020 U.S. dollars.
- For interpretability and to address negative values, income is transformed into the logarithmic form:

inctot → ln(inctot)

Independent Variable

- Following the 'Industry Codes by 2022 Census classification' IPUMS framework, individuals' primary industry is coded 0-9920.
- The technical consulting services sector (code: 7390) has been converted into a dummy variable indicator (Consulting = 1, otherwise 0).

ind → ind_new = ind_n

Control Variables

- ***Age:*** 18-64, continuous.

- ***Sex:*** Dummy variable for gender (1 if male, 0 if female).

- ***Veteran Status:*** Dummy variable (1 if veteran, 0 otherwise).

- ***Marital Status:*** Dummy variable (1 if married, 0 otherwise).

- ***Family Income:*** Total annual family income earned by households in dollars. Transformed like personal income, see above. *Ftotinc → ln(ftotinc)*

- ***Metropolitan Status:*** Dummy variable (1 if in large city/metropolitan area, 0 otherwise)

- ***Education:*** Controlled for those with a bachelor's degree or higher.

---

3. Methodology

Data Cleaning and Preparation

- Handling missing values: Observations with missing income data were dropped, while categorical variables with missing responses were assigned "unknown" categories to preserve sample size.
- Variable transformation: To address skewness in income data, I log-transformed total annual earnings as well as total family income earnings.
- Industry classification: The dataset contained over 200 industry categories. To enhance interpretability, I consolidated these into six broad industry groups based on sector similarities and transformed into a dummy variable.
- Dummy variable creation: I generated binary factors for key demographics such as gender, education level, and veteran status.

To estimate the relationship between industry and income, I use the following log-linear regression model:

$$ln(inctot) = \beta_0 + \beta_1 ind\_n + \beta_2 age + \beta_3\ educ + \beta_4 gender + \beta_5 vetstatus + \beta_6 maritalstatus + \beta_7 ln(ftotinc) + \beta_8 metropolitan\ status + u_i$$

where u_i is the error term capturing unobserved factors affecting income.

The coefficient of interest is β1, representing the percentage difference in income between technical consulting employees, holding other factors constant.

All regressions are estimated using **ordinary least squares (OLS)** with robust standard errors.

---

4. Results

Descriptive Statistics

Individuals in the sample report a mean annual income of approximately $102,000 compared to $239,000 for family income. The gender difference was a near even split, with males proportionately accounting for 48% of the sample and 52% for females. About 60% of the sample reported just 4 years of post-secondary education, whereas 40% reported 5+ years of post-secondary education.

Regression Estimates

The regression results indicate a statistically significant and economically meaningful beta coefficient:

Figure 1: Effect of Industry on Log of Personal Income

Variable	Industry	Industry & Gender	Industry, Gender, & Occupation	Industry & All Else
Industry(β_1)	*.2948*	*.2678*	*.0833*	*.1193*
	(.0088)	*(.0086)*	*(.0084)*	*(.0060)*
Robust SE's	*(.0084)*	*(.0083)*	*(.0082)*	*(.0059)*
T-Stat w/ RSE's	*35.095*	*32.265*	*10.159*	*20.220*
Controls	None	Gender	Gender & Occupation	All Else
Adjusted R^2	*.0020*	*.0456*	*.1057*	*.5634*

Notes:

Non-robust standard errors are provided immediately below the coefficient estimates and are in parentheses. The robust SE's are provided after conducting the B-P and White heteroskedasticity tests. The 'all else' variables are listed in section 2. Outliers were further controlled for by dropping any observations from the dataset if the total personal income level was less than 100. This reduction slightly affected the coefficients and SE's in the table above.

Interpretation

The coefficient on the technical consulting industry (*.1193*) suggests that, holding other factors constant, individuals working in consulting earn approximately 11.93% higher income, on average, than comparable individuals in other industries.

As expected, education, *age,* and *all else* have positive effects on income, reflecting returns to human capital and work experience. The additional factors are significant to include in the model as they increased the beta coefficient factor from *.0833* to *.1193*.

Gender remains a strong predictor, still positively impacting the consulting industry's effect on annual income.

Occupation within industry is also statistically significant. When occupation controls are introduced, the industry coefficient decreases substantially from *.2678* to *.0833*, suggesting that a portion of the initial consulting coefficient reflects occupational composition – that is, consulting employs a greater share of high-paying occupations relative to other industries.

---

5. Discussion

The statistical significance derived from the T-Statistic and p-values likely reflect a combination of skill-based pay differentials, performance-driven compensation, and selection effects—as individuals entering consulting may already possess stronger analytical and communication abilities.

However, the findings should be interpreted with caution due to potential omitted variable bias. Variables such as geographic region, hours worked, and firm size could further explain income variation across industries. Additionally, the cross-functionality of practically every industry and occupation severely limits causal inference.

Despite these limitations, the results align with broader evidence on wage disparity across industries and underscore the strong labor-market performance of technical consulting professionals.

---

6. Conclusion

This analysis finds that employment in the consulting industry is associated with approximately 11.93% higher income, controlling for education, age, gender, and other demographic factors.

The results reinforce the idea that industry choice plays a significant role in shaping individual income trajectories. Future research could extend this work by examining the effects over a greater time period or emphasizing how they differ across firm sizes and regions.

References

- Mincer, J. (1974). *Schooling, Experience, and Earnings.* NBER.

- Team, MPC UX/UI. "U.S. Census Data for Social, Economic, and Health Research." *IPUMS USA*, NIH, 2022.

- Wooldridge, J. M. (2019). *Introductory Econometrics: A Modern Approach.* 7th Edition.