

## History 595: Tenth Assignment

Tenth Assignment, due Wednesday, May 9, 2:00 PM in electronic form to [margo@uwm.edu](mailto:margo@uwm.edu) and on paper in class.

The Cost of Living dataset we have from the iron and steel workers in Pennsylvania in 1889-90 allows us to look into the day to day lives of working class Americans in the Gilded Age. One question we can ask is whether these families were in debt or were earning more than they spent, i.e., whether they had surplus income to be put to savings. We know what their total household income was, and we also know what they spent on all their expenditures. The difference between their income and expenditures is a measure of surplus or debt. We are interested in understanding if they were in debt or surplus and then what the determinants of the situation are. We'll do this by explore the basic patterns and then testing regression models to understand the situation further.

Variables:

Savings: Total Income - Total Expenditures in \$

Husband's Income (v33): in \$

Income from Children (v35): in \$

Union: 1 - Expenditure for unions;

0 = No expenditures for unions

US: 1= North American nationality;

0= Not North American nationality

Own: 1= owns home;

0 = doesn't own home

Liquor: 1= Household has expenditures for alcoholic beverages;

0 = Household does not have expenditures for alcoholic beverages

Sickness and Death (v96): \$ amount the household spent on "sickness and death"

Total in Family (v12)

Answer the questions from the information in the output.

1. Are the households on average showing a surplus or a deficit compared to their expenditures?

Report the statistic you are using to support your conclusion.

2. Describe the dispersion in the distribution of the savings variable. Report the statistics you are using to support your analysis and explain why you report the information you did.

I've limited the analyses to cases where the husbands makes less than \$1,000 per year. Just using a bivariate regression model with savings as the dependent variable and husband's income as the independent variable, the regression equation is:

$$\text{Savings} = -130.16 + .325 * \text{Husband's income.}$$

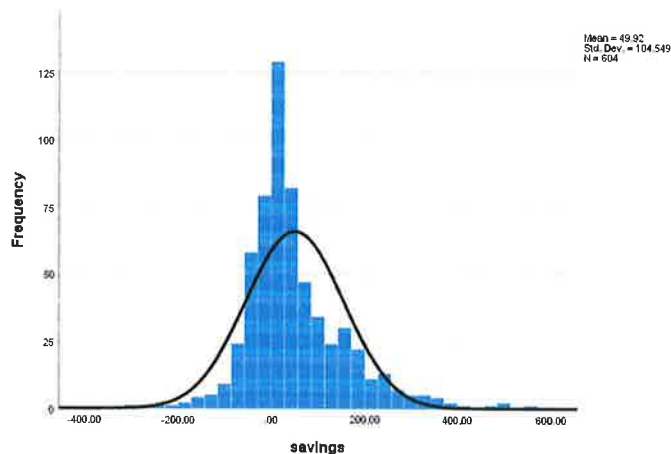
The equation is statistically significant and accounts for 23% of the variance in the distribution of savings.

The next pages of output are reports from a multivariate regression model.

See the output from a multiple regression model with the additional variables (8 in total): Husband's Income (v33); Income from Children (v35); Union; US; Own; Liquor; Sicknes and Death (v96); Total in Family (v12)

Answer the following questions.

3. Report the proportion of households with expenditures for alcoholic beverages
4. Report the proportion of households with North American nationality
5. Report the mean income in these households from children's work.
6. Report the proportion of households that own their home.
7. Write the regression equation from the multivariate model.
8. Is the model statistically significant? How do you know?
9. What proportion of the variance in savings in these households is accounted for by the eight variables in the model?
10. Interpret the model. Explain to someone who has not taken History 595 what the model shows about the patterns determining a household's savings or deficit.



# Regression

## Descriptive Statistics

	Mean	Std. Deviation	N
savings	49.8604	104.62729	603
HUSBAND-S INCOME	553.4316	154.74261	603
liquor	.4229	.49443	603
union	.3367	.47296	603
CHILDREN-S INCOME	43.7715	149.30436	603
SICKNESS AND DEATH	16.1964	22.98044	603
TOTAL IN FAMILY	4.85	2.050	603
own	.1990	.39958	603
us	.5987	.49057	603

## Variables Entered/Removed<sup>a</sup>

Model	Variables Entered	Variables Removed	Method
1	us, union, SICKNESS AND DEATH, CHILDREN-S INCOME, own, liquor, HUSBAND-S INCOME, TOTAL IN FAMILY <sup>b</sup>		Enter

a. Dependent Variable: savings

b. All requested variables entered.

## Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.644 <sup>a</sup>	.415	.407	80.53663

a. Predictors: (Constant), us, union, SICKNESS AND DEATH, CHILDREN-S INCOME, own, liquor, HUSBAND-S INCOME, TOTAL IN FAMILY

### ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2737244.028	8	342155.504	52.752	.000 <sup>b</sup>
	Residual	3852772.006	594	6486.148		
	Total	6590016.034	602			

a. Dependent Variable: savings

b. Predictors: (Constant), us, union, SICKNESS AND DEATH, CHILDREN-S INCOME, own, liquor, HUSBAND-S INCOME, TOTAL IN FAMILY

### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-112.289	15.868		-7.077	.000
	HUSBAND-S INCOME	.365	.023	.540	16.165	.000
	liquor	.820	6.980	.004	.117	.907
	union	-39.861	7.570	-.180	-5.266	.000
	CHILDREN-S INCOME	.191	.024	.273	7.983	.000
	SICKNESS AND DEATH	-.870	.146	-.191	-5.955	.000
	TOTAL IN FAMILY	-8.666	1.790	-.170	-4.841	.000
	own	61.408	8.561	.235	7.173	.000
	us	14.405	6.998	.068	2.058	.040

a. Dependent Variable: savings