

Women in the Workforce: A Wonderful Addition or a Woeful Mistake?

EC 315

Sample Report

Term Project Report

Fall I 2007

TABLE OF CONTENTS

BACKGROUND	3
DISCUSSION OF RESULTS	4
SUMMARY	8
REFERENCES	9
APPENDIX	10

Women in the Workforce: A Wonderful Addition or a Woeful Mistake?

Background

Many human resources professionals, scholars, feminists, and economists tout the addition of women to the U.S. workforce. Wendell French (2005) speculates in *Human Resources Management* that the continuous stream of women entering the workforce will explain a 55% increase in total U.S. labor force expansion between the years of 2002 and 2012 (p. 57). In addition, the percentage of working women continues to increase (French, 2005, p. 57). As women comprise an increasingly larger share of the labor market, their contributions, education, and effect on the economy warrants discussion. The aim of this project is to determine the effects of the entrance of larger proportions of increasingly educated women over the age of 25 with 4 years of college on the productivity of non-farm business in the United States, while holding the rate of population growth for this specific class (females over age 25), average number of hours worked, and average salary constant. This study employs a time-series analysis with observations from 1966 to 2006 included. Demographic data on education was taken from the U.S. Census Bureau and productivity information from the Bureau of Labor Statistics. The model (less constants and coefficients is):

$$\text{OUTPUT} = \% \text{COLLEGE_FEM} + \text{AVG_SAL} + \text{AVG_HOURS} + \text{POP_GROWTH}$$

The result or dependent variable, OUTPUT, includes non-farm, seasonally adjusted output per hour. This variable is calculated using the ratio of the output of goods and services to labor hours required to produce them. %COLLEGE_FEM, the first independent variable, is the percentage of the female population age 25 and over who have completed 4 years of college and is compiled by the U.S. Census Bureau. This measure is used because of the established relationship between productivity and higher learning. If other independent variables are held

constant, increases in education should result in a positive change in productivity (Sweetman, 2002 & Saxton, 2000). AVG_SAL is the real hourly compensation received by employees in non-farming business sectors. It is seasonally adjusted and indexed to 1992. This figure is utilized in the formula because wages have become a progressively significant incentive for workers to remain in or become more productive in the workforce. This being said, a positive relationship between average wages and productivity should exist (Fazzari, 2007). The average weekly hours spent on the job is another possible predictor of the dependent variable, $OUTPUT$. The Census Bureau collects this data from American workers for the Bureau of Labor Statistics. Many employers make adjustments to the hours which employees work in order to affect changes in productivity (International Labour Office Geneva, 2007). The additional time spent on the job increases output; essentially this should mean that a positive relationship exists between average hours and productivity, all other independent variables held equal (Skocnylas & Tissot 2004). Finally, as the population grows so does potential, equilibrium, and per capita output; this should also affect hourly productivity in a positive fashion (Fazzari, 2007).

Discussion of Results

The model was regressed and yielded the following results:

Regression equation is: $OUTPUT = -100 + 1.50 \%COLLEGE_FEM + 1.27 AVG_SAL + 0.497 AVG_HOURS + 4.46 POP_GROWTH$

DEPENDENT VARIABLE: OUTPUT		ADJUSTED R ² = .9914	n = 41
Independent Variables	Coefficient	Student t	Significance of t
%COLLEGE_FEM	1.500940684	5.820633573	1.20689E-06
AVG_SAL	1.268280718	12.52026283	1.1126E-14
AVG_HOURS	0.49737168	3.052832861	0.004
POP_GROWTH	4.461596032	1.187963082	0.24262667
Durbin Watson = .802773			

The focus of this analysis is on the impact of %COLLEGE_FEM on OUTPUT. As evident above, the Durbin Watson Statistic of .802773 fell into the rejection region, indicating positive autocorrelation. Autocorrelation occurs when a pattern exists between the error terms due to a variable missing from the analysis. In this regression, the coefficients of the independent variables are biased to an unknown extent and are not reliable or reportable in a scholarly paper, publication, or report.

If the Durbin-Watson for this regression had passed, the independent variables %COLLEGE_FEM and AVG_SAL would have been significant at $\alpha = .05$, .025, .01, .005, and .001. The variable AVG_HOURS would have shown significance at $\alpha = .05$, .025, .01, and .005. However, the independent variable, POP_GROWTH shows inadequate significance with a p value of .2426; this does not meet the criteria of showing significance at the $\alpha = .05$, or even .10 level; however, it is significant when the sig value is less than or equal to α .

The R^2 value of 99.2% suggests that the independent variables account for 99.2% of the variation of the outcome; this is often the case in time series regressions, in which one observation builds upon another. R^2 cannot be relied upon since the Durbin-Watson indicates positive autocorrelation. It is desirable for R^2 to be 50% or greater. The Adj. R^2 value of 99.1% is also thought of as a "good" thing, if the Durbin-Watson had passed. In the case of adjusted R^2 , this regression indicates that the independent variables explain 99.1% of the variance of the dependent variable. Both statistics must be less than or equal to one, but greater than or equal to 0. Often, people rely heavily on the use of R^2 and Adj. R^2 , disregarding the Durbin-Watson test. Since this analysis is a time series regression, the Durbin-Watson is more valuable to the analyst than the R^2 and Adj. R^2 . Without a passing Durbin-Watson test for a time-series analysis, both of the preceding are useless, as is the case with this regression.

Another important consideration when regressing an equation is the presence of multicollinearity. In this model, there was a complete absence of it. All pairs of independent variables were regressed, and the resulting R^2 from the bivariate regressions compared to the R^2 of the entire model. The results are detailed below:

Regression Statistics % College Females & Avg Salary

Multiple R	0.960929531
R Square	0.923385564

Regression Statistics % College Fem & Avg Hours

Multiple R	0.82176622
R Square	0.67529972

Regression Statistics for % College Fem & Pop Growth

Multiple R	0.245677796
R Square	0.060357579

Regression Statistics for Avg_Sal & Avg_Hours

Multiple R	0.865413869
R Square	0.748941165

Regression Statistics Avg_Sal & Pop_Growth

Multiple R	0.21991564
R Square	0.048362889

Regression Statistics of Avg_Hours & Pop_Growth

Multiple R	0.348112319
R Square	0.121182187

As is evident above, the bivariate regressions yield R^2 values less than the value of the entire regression. Multicollinearity happens when two more of the predictors have a linear relationship. When this occurs, the statistical software package does not know which variable to give the coefficient to. Often, one coefficient will be near zero while the other coefficient of the collinear variable will be the source of all the affect on the outcome, causing the coefficients to be biased. The absence of multicollinearity indicates that the coefficients are not biased due to its presence.

Another useful way to analyze the effect of independent variables on the outcomes is through coefficients. The coefficients for the model are detailed below:

Independent Variables	Coefficient
%COLLEGE_FEM	1.500940684
AVG_SAL	1.268280718
AVG_HOURS	0.49737168

POP_GROWTH	4.461596032
Durbin Watson = .592049	

The 1.5009 coefficient value for %COLLEGE_FEM indicates that for every one percent increase in the percentage of females who have obtained four years of college, output increases by 1.5009. 1.2682 is the coefficient for the predictor, AVG_SAL and indicates that for every single unit increase in salary, output increases by 1.268. The .4973 coefficient value for AVG_HOURS suggests that for 1 hour increase, output is increased by .4973. POP_GROWTH brings with it a coefficient of 4.4615 indicating that for every 1,000 people that are added to the population, output increases by 4.4616. While these coefficients indicate a fairly significant relationship there are two things that must first be considered:

- The Durbin-Watson test failed, which means that the coefficients of the predictors are biased to an unknown extent.
- In the case of POP_GROWTH, the p value method of hypothesis test indicates that the variable is insignificant.

Had the Durbin-Watson passed, and the sig value of POP_GROWTH been less than or equal to alpha, the coefficients of the variables would yield the results detailed above. According to the coefficients, population growth has the most significant effect on the outcome, followed by the percentage of women with 4 years of college, average salary, and finally average hours.

Summary

Unfortunately, the failure of the Durbin-Watson test makes the model biased to an unknown extent. Unless the missing variable can be found, the results are essentially useless. If the Durbin-Watson had passed, all variables except POP_GROWTH would be significant predictors of the outcome. This would indicate that as the number of women who have 4 years of college increases, so does output per hour. In addition, the predictions made in the Background section would prove to be true.

Appendix

Year	OUPUT	%COLLEGE_FEM	AVG_SAL	AVG_HOURS	POP_GROWTH
1966	63.585	5.3	73.197	113.286	1.16
1967	64.687	5.6	75.15	111.415	1.09
1968	66.893	6.6	77.778	110.788	1
1969	66.993	5.8	78.773	110.013	0.98
1970	67.988	5.8	79.842	108.212	1.17
1971	70.703	5.9	81.365	107.678	1.26
1972	73.061	6.3	83.953	107.885	1.07
1973	75.336	6.6	85.479	107.506	0.95
1974	74.208	6.8	84.493	105.93	0.91
1975	76.221	7.1	85.213	104.435	0.99
1976	78.728	7.4	87.37	104.562	0.95
1977	79.984	8	88.715	103.956	1.01
1978	81.022	7.9	90.25	103.618	1.06
1979	80.739	8.1	90.249	102.934	1.1
1980	80.579	8.9	89.999	101.779	0.96
1981	81.691	8.6	90.211	101.381	0.98
1982	80.831	8.62	91.107	100.764	0.95
1983	84.458	9.2665	91.109	101.696	0.91
1984	86.121	9.7429	91.121	102.403	0.87
1985	87.456	10.0016	92.113	102.189	0.89
1986	90.151	10.0983	95.175	101.198	0.92
1987	90.608	12.7948	95.523	101.429	0.89
1988	92.106	10.6	96.675	101.118	0.91
1989	92.783	11.1578	95.064	101.612	0.94
1990	94.515	11.48	96.047	100.465	1.07
1991	96.06	11.86	97.414	99.679	1.08
1992	100	13.1	100	100	1.14
1993	100.411	13.4596	99.478	100.556	1.08
1994	101.524	13.67	99.061	100.927	0.99
1995	102.009	14.0052	98.773	100.832	0.95
1996	104.715	15.1403	99.45	100.257	0.92
1997	106.415	15.3831	100.352	100.955	0.96
1998	109.354	16.95	104.89	100.867	0.92
1999	112.508	16.1	107.542	101.268	0.9
2000	115.689	16.3	111.56	100.514	1.01
2001	118.583	16.8	112.848	99.17	1.0102
2002	123.473	17.1915	115.098	98.947	1.0102
2003	128.034	17.458	117.076	98.505	1.00927
2004	131.542	17.606	118.166	98.51	1.00977
2005	134.097	17.8933	118.94	98.322	1.0098

References

- International Labour Office; Geneva, (2007). Working time around the world: Main findings and policy implications. Retrieved August 29, 2007, from International Labour Office Web site: http://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/documents/publication/wcms_082838.pdf
- Fazzari, (2007, April 17). Retrieved September 2, 2007, from Washington State University, St. Louis Web site: artsci.wustl.edu/~ec104sf/Lec%20Notes%20104-8.doc
- French, W.L. (2005). *Human Resources Management*. New York: Houghton Mifflin Company.
- Saxton, Jim (January 2000). Joint Economic Committee Study. Retrieved September 1, 2007, from The United States House of Representatives Web site: <http://www.house.gov/jec/educ.htm>
- Skoczylas, L., & B. Tissot (2005). Revisiting Recent Productivity Developments Across OECD Countries. *Bank for International Settlements*, Retrieved September 2, 2007, from <http://www.bis.org/ifccommittee.org/tissot.pdf>.
- Sweetman, A. (2002, November 27). Working smarter: Education and productivity. *The Review of Economic Performance and Social Progress*, Retrieved September 1, 2007, from <http://www.irpp.org/miscpubs/archive/rep1202/sweetman.pdf>

2006	135.393	18.1343	119.664	98.555	1.00975
------	---------	---------	---------	--------	---------

Minitab Regression Calculation

Worksheet size: 10000 cells.

Welcome to Minitab, press F1 for help.

Regression Analysis: OUPUT versus %COLLEGE_FEM, AVG_SAL, ...

The regression equation is
 OUPUT = - 100 + 1.50 %COLLEGE_FEM + 1.27 AVG_SAL + 0.497 AVG_HOURS
 + 4.46 POP_GROWTH

Predictor	Coef	SE Coef	T	P
Constant	-100.40	21.17	-4.74	0.000
%COLLEGE_FEM	1.5009	0.2579	5.82	0.000
AVG_SAL	1.2683	0.1013	12.52	0.000
AVG_HOURS	0.4974	0.1629	3.05	0.004
POP_GROWTH	4.462	3.756	1.19	0.243

S = 1.87253 R-Sq = 99.2% R-Sq(adj) = 99.1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	16156.6	4039.2	1151.94	0.000
Residual Error	36	126.2	3.5		
Total	40	16282.8			

Source	DF	Seq SS
%COLLEGE_FEM	1	15517.5
AVG_SAL	1	587.1
AVG_HOURS	1	47.1
POP_GROWTH	1	4.9

Unusual Observations

Obs	%COLLEGE_FEM	OUPUT	Fit	SE Fit	Residual	St Resid
22	12.8	90.608	94.368	0.617	-3.760	-2.13R
35	16.3	115.689	120.049	0.587	-4.360	-2.45R

R denotes an observation with a large standardized residual.

Durbin-Watson statistic = 0.802773

$D_1 = 1.29$, which is greater than .802773, signaling the failure of the DW test.

Test for Multicollinearity

SUMMARY OUTPUT

<i>Regression Statistics % College Females & Avg Salary</i>	
Multiple R	0.96092953
R Square	0.92338556
Adjusted R Square	0.92142109
Standard Error	1.18075699
Observations	41

<i>Regression Statistics % College Fem & Avg Hours</i>	
Multiple R	0.82176622
R Square	0.67529972
Adjusted R Square	0.66697407
Standard Error	2.43078502
Observations	41

<i>Regression Statistics for % College Fem & Pop Growth</i>	
Multiple R	0.2456778
R Square	0.06035758
Adjusted R Square	0.03626418
Standard Error	4.13510486
Observations	41

SUMMARY OUTPUT

<i>Regression Statistics for Avg_Sal & Avg_Hours</i>	
Multiple R	0.86541387
R Square	0.74894117
Adjusted R Square	0.74250376
Standard Error	6.26476959
Observations	41

<i>Regression Statistics Avg_Sal & Pop_Growth</i>	
Multiple R	0.21991564
R Square	0.04836289
Adjusted R Square	0.02396194
Standard Error	12.1970003
Observations	41

Regression Statistics of Avg_Hours & Pop_Growth

Multiple R	0.34811232
R Square	0.12118219
Adjusted R Square	0.0986484
Standard Error	3.66430833
Observations	41