Brendan Dennehy Case Scenario

Brendan Dennehy works for Transon Investments, PLC, a Dublin-based hedge fund with significant equity investments in technology companies in Asia, North America, and Europe. Transon is concerned about the recent poor performance of one of the fund's Chinese investments, Winston Communications, an assembler of telecommunications equipment. Transon's chief of information technology (IT) is Sean Malloy.

Yesterday, Winston's IT office sent Malloy data related to the assembly process and a printout of an analysis of the number of defective assemblies per hour. Winston's IT people believe that the number of defective assemblies per hour is a function of the outside air temperature and the speed (production rate) of the assembly lines. Malloy recalls that Dennehy has had substantial training in statistics while working on his MBA. He asks Dennehy to help him interpret the regression results supplied by Winston

EXHIBIT 1 REGRESSION RESULTS

$D_t = b_0 + b_1\text{Air}_t + b_2R_t + e_t$

	Coefficient	Standard Error
Constant (b_0)	0.016	0.0942
Outside air temperature (b_1)	0.0006	0.001
Assembly line speed (b_2)	0.5984	0.3
Number of observations used in the regression		384
Critical t-value at 5% significance (two-tail test where coefficient equals zero)		1.96

R^2	Standard Error of the Estimate	Durbin–Watson Statistic	F-Statistic	Significance of F
0.414	0.333	1.89	157.699	0
Durbin–Watson critical values (5% significance)		1.63		1.72
Correlation between outside air temperature and assembly line speed			0.015	

Using the data provided in Exhibit 1, Dennehy tests the hypothesis that the coefficients for outside air temperature and assembly line speed are significantly different from zero, using a significance level of 5%. Dennehy also uses the results given in Exhibit 1 to evaluate the potential for multicollinearity in the data.

Finally, Dennehy would like to confirm that nonstationarity is not a problem. To test for this he conducts Dickey–Fuller tests for a unit root on each of the time series. The results are reported in Exhibit 2.

EXHIBIT 2 RESULTS OF THE DICKEY-FULLER TESTS

Time Series	Value of the Test Statistic	Standard Error	t-Statistic	Significance of t
Defective assemblies per hour	0.0036	0.0023	1.591	0.1123
Outside air temperature	–0.423	0.0724	–5.846	0
Assembly line speed	–0.586	0.043	–13.510	0

Dennehy tells Malloy about the Dickey-Fuller test results, stating:

"We can safely use regression to estimate the relationship between the dependent variable and the independent variables if 1) none of the three time series exhibit a unit root or 2) all three time series exhibit a unit root but they are also mutually cointegrated."

1Q. Based on Exhibit 1 and statistical tests, the best conclusion Dennehy can make is that the regression coefficient is significantly different from zero with respect to the coefficient(s) for:

A. assembly line speed (b_2) only.
B. both outside air temperature (b_1) and assembly line speed (b_2).
C. outside air temperature (b_1) only.

2Q. The most appropriate interpretation of the results reported in Exhibit 1 is that:

A. the F-statistic of the regression is not significant.
B. predictions of defective assemblies per hour made using the regression have only about a 41% chance of being correct.
C. variations in the independent variables explain approximately 41% of the variation in the defective assemblies per hour.

3Q. What is the most appropriate inference from the Durbin–Watson statistic reported in Exhibit 1? The Durbin–Watson test:

A. is inconclusive.
B. rejects the null hypothesis of no positive serial correlation.
C. fails to reject the null hypothesis of no positive serial correlation.

4Q. The results reported in Exhibit 1 are most accurately interpreted as indicating that:

A. the reported R^2 is spurious.

B. multicollinearity is not present.

C. the regression coefficients have inflated standard errors.

5Q. Assuming a 5% level of significance, the most appropriate conclusion that can be drawn from the Dickey–Fuller results reported in Exhibit 2 is that the:

A. test for a unit root is inconclusive for the dependent variable.

B. independent variables exhibit unit roots but the dependent variable does not.

C. dependent variable exhibits a unit root but the independent variables do not.

6Q. Dennehy's statement about the Dickey–Fuller test is best characterized as:

A. incorrect because only the independent variables series need to be tested for the absence of a unit root

B. incorrect because only the dependent variable series needs to be tested for the absence of a unit root.

C. correct.

Solution

1：**A is correct.** The null hypotheses are that the coefficients equal zero. The alternative hypotheses are that the coefficients do not equal zero (two-tailed tests). The appropriate test statistics, t, are calculated by dividing the estimates of the coefficients by their respective standard error.

t_{b1} = 0.0006/0.0010 = 0.60

t_{b2} = 0.5984/0.30 = 1.9947

The test statistic for outside air temperature is less than the critical value of 1.96. The test statistic for assembly line speed exceeds the critical value of 1.96. Dennehy cannot reject the null hypothesis that the population regression coefficient for outside air temperature, b_1, is zero. Dennehy can reject the null hypothesis that b_2 is zero at the 5% level of significance.

B is incorrect because only assembly line speed is significant.

C is incorrect because outside air temperature is not significant.

2：**C is correct.** The R^2 indicates that variations in the independent variables explain approximately 41% of the variation in the dependent variable.
The F-statistic is highly significant. R^2 does not inform us regarding the probability of a dependent variable prediction being correct.

A is incorrect because the F-test of the regression is highly significant.

B is incorrect because R^2 does not inform us regarding the probability of a dependent variable prediction being correct.

3：**C is correct.** The value of the Durbin–Watson statistic is given in Exhibit 1 as 1.890. The critical values are given as 1.63 and 1.72. Because the value (1.890) exceeds the upper critical value (1.72), the Durbin–Watson test fails to reject the null hypothesis of no positive serial correlation.

A is incorrect because a value between the lower Durbin–Watson and the upper Durbin–Watson is inconclusive.

B is incorrect because rejection of the null requires a Durbin–Watson below the lower critical value.

4：**B is correct.** The pairwise correlation is low. The only case in which correlation between independent variables may be a reasonable indicator of multicollinearity occurs in a regression with exactly two independent variables, as is the case in this problem. Furthermore, the additional classic symptoms of multicollinearity (high R^2 and significant F-statistic but not significant coefficients) are not present.

A is incorrect because even in the face of multicollinearity, a regression may have a high R^2.

C is incorrect because at least one coefficient (b_2) is different from zero.

5：C is correct. The Dickey–Fuller test uses the following type of regression:

$$x_t - x_{t-1} = b_0 + g_1 x_{t-1} + \varepsilon_t, E(\varepsilon_t) = 0$$

The null hypothesis is H_0: $g_1 = 0$ versus the alternative hypothesis H_a: $g_1 < 0$ (a one-tail test). If $g_1 = 0$, the time series has a unit root and is nonstationary. Thus, if the null hypothesis fails to be rejected, then the possibility exists that the time series has a unit root and is nonstationary.

Based on the t ratios and their significance levels in Exhibit 2, the null hypothesis that the coefficient is zero is rejected for both outside air temperature and assembly line speed (i.e., the independent variables). But the null hypothesis is not rejected for the dependent variable, defective assemblies per hour.

A is incorrect because the p-value for the Dickey–Fuller test of the dependent variable time series (11.23%) clearly fails to reject the null at 5% level of significance.

B is incorrect because we reject the null hypothesis that the coefficient is zero for both outside air temperature and assembly line speed. We do not reject the null for the independent variable, defective assemblies per hour.

6：C is correct. One possibility is that none of the time series used in a regression exhibit a unit root. In that case, regression analysis can safely be used. Alternatively, if at least one time series (the dependent variable or one of the independent variables) has a unit root while at least one time series (the dependent variable or one of the independent variables) does not, the error term in the regression cannot be covariance stationary. Consequently, multiple linear regression should not be used to analyze the relationship among the time series in this scenario. Another possibility is that each time series, including the dependent variable and each of the independent variables, has a unit root. If this is the case, it needs to be established whether the time series are cointegrated. When all series used in a regression display unit roots, but they are also mutually cointegrated, regression analysis can safely be used. A and B are incorrect.