Econometrics Basics: Dealing with Simultaneity Bias

John E. Floyd
University of Toronto
July 22, 2013

This document outlines the nature of simultaneity bias in regression analysis and the standard method for dealing with it, two-stage least squares. A classic example, which we begin with, is the estimation of supply and demand curves where the two important variables of interest, price and quantity, are simultaneously determined. We will then apply both economic theory and econometrics to an analysis of the simultaneity bias issues involved in understanding the determination of real exchange rates under conditions of international capital mobility. Indeed, at that point the focus becomes as much on economic theory as on econometrics.

Two-Stage Least Squares: A Simple Example

Figure 1 on the next page presents a situation where the equilibrium price and quantity are determined by the intersection of the demand and supply curve and the supply curve has shifted at two points in time, once upward and once downward, to new levels but the demand curve has remained unchanged. As the price falls the quantity increases along the demand curve in a manner consistent with the demand function

\[ Q = \alpha + \beta P \]  \hspace{1cm} (1)

representing a perfect fit with \( \alpha > 0 \) and \( \beta < 0 \).

---

\(^1\)An excellent more detailed analysis of the material covered here can be found in James H. Stock and Mark W. Watson, *Introduction to Econometrics*, Addison Wesley, 2003, Chapter 10.
Suppose now that, as shown in Figure 2 above, the demand curve has also shifted a couple of times, downward in the period when the supply curve shifted to the left and upward in the period when it shifted to the right. Were we to regress the quantity on the price, we would be trying to estimate the population equation represented by the line DD

\[ Q_t = \alpha + \beta P_t + \epsilon_t \]  

(2)

where \( \epsilon_t \) represents an error term that is negatively correlated with \( P_t \). But the actual regression equation would be

\[ Q_t = \hat{\alpha} + \hat{\beta} P_t + u_t \]  

(3)

where \( \hat{\alpha} > \alpha \) and \( |\hat{\beta}| > |\beta| \) and, given that we have only three observations, \( u_t \to 0 \) for all \( t \). The estimated demand curve will be a line joining the points a, b and c in Figure 2. Notice here that we are regressing the element along the horizontal axis (quantity) on the element along the vertical axis (price), so the slope of the estimated DD line is actually \( 1/\hat{\beta} \) and the error associated with each observation, given by \( u_t \), is the difference between the actual quantity and the quantity indicated by the estimated DD curve.

Suppose alternatively that the upward and downward shifts of the demand curve occurred in the same periods as the respective upward and downward shifts of the supply curve as shown in Figure 3 above. The population demand curve (2) remains the same except the error term \( \epsilon_t \) is now positively correlated with \( P_t \). And our regression equation now becomes
\[ Q_t = \tilde{\alpha} + \tilde{\beta} P_t + u_t \]  

where \( \tilde{\alpha} < \alpha \) and \( |\tilde{\beta}| < |\beta| \) and, as before, \( u_t \to 0 \) for all \( t \). The estimated demand curve will now be a line joining the points \( a, b \) and \( c \) in Figure 3. This estimated demand curve will now be steeper than the true one whereas the estimated curve in Figure 2 was flatter than the true one.

It should now be clear that if the demand curve were to shift by greater amounts than the supply curve, the estimated regression line would have the wrong slope—it would become a poor estimate of the supply curve SS.

Clearly, the fundamental problem involved in the estimation of the demand and supply curves is that the error term \( \epsilon_t \) is correlated with the price of the good \( P_t \). Of course, in a situation where there were many periods involved that error term will take many values. And if the variance of it happened to be small with the shifts in the supply curve on average many times greater than the shifts of the demand curve, our regression equation would give us quite accurate measures of \( \alpha \) and \( \beta \) when the sample size is sufficiently large.

A completely opposite situation will arise when the shifts in the demand curve are substantially greater than the shifts in the supply curve, as shown in Figure 4 below. Imagine that the supply and demand curves fall within the ranges denoted over many periods so that the equilibrium points form a dense mass within the area \( a b c d \). In this case our OLS regression of quantity on price will estimate the supply curve

\[ Q_t = \delta + \gamma P_t + \vartheta_t \]  

with our estimate of \( \delta \) being negative and slightly smaller than its true value and our estimate of \( \gamma \) being positive and slightly larger than its true value. And as the range in which the supply curve shifts gets smaller relative to the range in which the demand curve shifts, our estimated values will become closer to the true ones. And, of course, as the range in which the supply curve shifts gets larger our estimates become increasingly inaccurate.

The problem is that the error terms \( \epsilon \) and \( \vartheta \) in the demand and supply curves are correlated with the price variables in the respective equations, biasing our estimates. These errors can usually be substantially reduced by adding the exogenous variables that shift the demand curve to the demand
equation and the exogenous variables that shift the supply curve to the supply equation. For example the level of income and the prices of substitute goods can be added as independent variables in the demand equation and the level of productivity and the prices of other goods that use the same resources can be added as independent variables in the supply equation. Unless these added exogenous variables can encompass everything determining the respective quantities demanded and supplied, however, correlations of the respective error terms with the price will remain.

The ultimate solution in this case is to find variables that are correlated with the price variable but not with the error terms in the respective equations. These are called \textit{instrumental variables}. The approach discussed here proceeds in two stages—and is therefore called \textit{two-stage least squares}. In the first stage we regress the price variable on the instrument or instruments for that equation as well as the other exogenous independent variables and save the fitted values. These fitted values of the price variable are uncorrelated with the error term in this first-stage regression due to the properties of regression analysis. Then we run a second stage regression of the quantity on the fitted values of the price variable, cleansed of correlation with the errors by the first-stage regression, plus the other independent variables affecting the demand or the supply according to which of the two
equations we are estimating. The resulting coefficient of the fitted value of the price level will, if the instruments are good ones, be an unbiased measure of the true coefficient of the price variable in the underlying demand or supply curve being estimated.

Let us proceed using an example presented by G. S. Maddala in his introductory econometrics textbook.\textsuperscript{2} He estimates the demand and supply curves of commercial banks’ loans to business firms in the United States using monthly data for the years 1979 through 1984. The variables are in Excel worksheet file loandata.xls, the Gretl data file loandata.gdt, the XLispStat data file loandata.lsp, and the tabular file loandata.tab which is readable by R. The variables are named for estimation purposes and described as follows.

\begin{itemize}
  \item QLOANS — Quantity of Commercial Loans Made by Banks
  \item PRIMRATE — Bank’s Prime Rate on Commercial Loans
  \item CBNDRATE — Interest Rate on Corporate Bonds
  \item TBRATE — 30-Day Treasury Bill Rate
  \item INDPROM — Industrial Production
  \item TOTDEP — Total Bank Deposits
\end{itemize}

Appropriate structural equations representing the demand and supply of commercial bank loans are:

\textbf{Demand}

\[ \text{QLOANS} = \beta_0 + \beta_1 \text{PRIMRATE} + \beta_2 \text{CBNDRATE} + \beta_3 \text{INDPROM} \]

\textbf{Supply}

\[ \text{QLOANS} = \delta_0 + \delta_1 \text{PRIMRATE} + \delta_2 \text{TBRATE} + \delta_3 \text{TOTDEP} \]

where $\beta_1$ and $\delta_2$ are negative and all the other coefficients are positive. This says that banks will expand their supply of loans in response to a higher prime rate, greater deposits, and a lower rate of return on alternative investments in treasury bills, and that commercial enterprises will increase their demand for loans in response to a fall in the prime rate, a rise in the cost of funding through corporate bond issues and an increase in their output.

Let us first fit the demand supply equations above in XLispStat using our OLSreg function. The appropriate code lines are as follows.

```
(load "ourfuncs")
(load "loandata")

(def regressand "Quantity of Loans")
(def regressors (list "Constant" "Prime Rate" "Corp. Bond Rate" "Ind. Prod.")
(def OLSDreg (OLSreg (bind-columns QLOANS) 
(bind-columns PRIMRATE CBNDRATE INDPROD) 1))

(def regressors (list "Constant" "Prime Rate" "T-Bill Rate" "Bank Deposits")
(def OLSSreg (OLSreg (bind-columns QLOANS) 
(bind-columns PRIMRATE TRATE TOTBDEP) 1))
```

And we obtain the results below for the demand and supply curves. All the XLispStat input code presented here, and more, is in the file simlbias.lsp and the output from that code is in the file simlbias.lou.

### ORDINARY LEAST SQUARES REGRESSION -- [DEMAND]

**Dependent Variable: Quantity of Loans**

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-stat</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-198.452</td>
<td>-2.861</td>
<td>0.003</td>
</tr>
<tr>
<td>Prime Rate</td>
<td>-15.923</td>
<td>-11.928</td>
<td>0.000</td>
</tr>
<tr>
<td>Corp. Bond Rate</td>
<td>35.915</td>
<td>14.147</td>
<td>0.000</td>
</tr>
<tr>
<td>Ind. Prod.</td>
<td>2.258</td>
<td>5.337</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Number of Observations:** 72  
**Degrees of Freedom:** 68  
**R-Squared:** 0.7799598762225266  
**Adjusted R-Squared:** 0.7702522237029321  
**Sum of Squared Errors:** 57135.44467779983  
**F-Statistic:** 80.34484904030259  
**P-Value** 0.0
## ORDINARY LEAST SQUARES REGRESSION

Dependent Variable: Quantity of Loans -- [SUPPLY]

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-stat</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-77.469</td>
<td>11.234</td>
<td>-6.896</td>
<td>0.000</td>
</tr>
<tr>
<td>Prime Rate</td>
<td>2.424</td>
<td>0.828</td>
<td>2.927</td>
<td>0.002</td>
</tr>
<tr>
<td>T-Bill Rate</td>
<td>-1.903</td>
<td>1.071</td>
<td>-1.777</td>
<td>0.040</td>
</tr>
<tr>
<td>Bank Deposits</td>
<td>0.332</td>
<td>0.006</td>
<td>51.318</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Number of Observations: 72  
Degrees of Freedom: 68  
R-Squared: 0.9767416046419779  
Adjusted R-Squared: 0.9757154989644181  
Sum of Squared Errors: 6039.256561301356  
F-Statistic: 951.891823622675  
P-Value 0.0

It turns out that the demand curve has a negative slope and the supply curve a positive one as our theory predicts. Because of the simultaneity involved, however, we can expect that these slopes and the other coefficients in the demand and supply equations will be biased.
The first step in running two-stage-least squares is to decide what the instrumental variables will be in the two equations. Here we follow Maddala, who notes that it would seem reasonable to expect that the two exogenous variables in the demand equation will be independent of the error term in the supply equation since they would be unlikely to shift the supply curve and the two exogenous variables in the supply equation will be independent of the error term in the demand equation in that they are unlikely to shift the demand curve. We thus use the exogenous variables in the demand equation as instruments for the supply equations, and the exogenous variables in the supply equation as instruments for the demand equation.

Since the first stage regression for each curve involves regressing the endogenous variable—the prime rate on business loans—on the instruments and exogenous variables for that equation, it turns out that the same first stage regression will be appropriate for both equations. The code for the first stage in XLispStat is thus

```lisp
(def regressand "Prime Rate")
(def regressors (list "Constant" "Corp. Bond Rate" "Ind. Prod." "T-Bill Rate" "Bank Deposits"))
(def TSTS1reg (OLSreg (bind-columns PRIMRATE) (bind-columns CBNDRATE INDPROD TBRATE TOTBDEP) 1))
; (def FITPRATE OLSfitted) ; Define the Fitted Values for Second Stage
```

The last line of code renames the series of fitted values for the prime rate on business loans as FITPRATE for use as an independent variable in both the second stage demand and supply regressions. The fitted values for this variable are uncorrelated with the error term in the first stage regression as a result of the nature of OLS estimation. The resulting first stage output is as follows.
ORDINARY LEAST SQUARES REGRESSION

Dependent Variable: Prime Rate

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-stat</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.998</td>
<td>3.014</td>
<td>0.663</td>
</tr>
<tr>
<td>Corp. Bond Rate</td>
<td>0.758</td>
<td>0.177</td>
<td>4.281</td>
</tr>
<tr>
<td>Ind. Prod.</td>
<td>0.007</td>
<td>0.022</td>
<td>0.314</td>
</tr>
<tr>
<td>T-Bill Rate</td>
<td>0.774</td>
<td>0.116</td>
<td>6.665</td>
</tr>
<tr>
<td>Bank Deposits</td>
<td>-0.005</td>
<td>0.002</td>
<td>-3.449</td>
</tr>
</tbody>
</table>

Number of Observations: 72
Degrees of Freedom: 67
R-Squared: 0.8682436231443801
Adjusted R-Squared: 0.8603775707947909
Sum of Squared Errors: 94.37302512953507
F-Statistic: 110.37857168465432
P-Value 0.0

The two instruments for the demand curve, the T-Bill Rate and Bank Deposits, are strong instruments—that is they are significant determinants of the endogenous variable PRIMRATE. One of the two instruments for the supply curve, industrial production, is clearly weak in that it is not a significant determinant of the prime loan rate. Anyway, we proceed along the same lines as Maddala does and use it as an instrument. On occasion, it may desirable to run F-tests for the significance of the instruments for the separate equations. To do this for the supply instruments, for example, we would run a restricted regression of the prime rate on the treasury bill rate and the quantity of bank deposits, dropping the other two variables. The sum of squared residuals from this regression minus the sum of squared residuals for the unrestricted regression above is a measure of the degree to which the variation of the residuals increases when we drop the two variables. The
resulting F-Statistic is

\[ F(n, df) = \frac{(RSSE - USSE)/n}{USSE/df} \]

where RSSE is the restricted sum of squared errors, USSE is the unrestricted sum of squared errors, \( n \) is the number of restrictions (2 in this case) and \( df \) is the degrees of freedom in the unrestricted regression. The parameters of the F-Statistic are thus \( n \) and \( df \), the number of degrees of freedom in the numerator and denominator respectively.\(^3\)

The second stage of two-stage least squares estimation also involves standard OLS regression but an important refinement is required. The standard errors of the coefficients produced by standard OLS turn out to be incorrect. The problem is that the standard OLS regression calculates the regression residuals as

\[ \hat{e} = y - F \hat{b} \]

where \( \hat{e} \) is the vector of second stage regression residuals, \( y \) is the vector of values of the dependent variable, \( F \) is the matrix of independent variables including a constant and the fitted values of endogenous variables from the first stage rather than the actual values and \( \hat{b} \) is the vector of coefficients. The calculated standard error of the regression, \( \hat{s}^2 \), is the sum of squares of these residuals divided by the degrees of freedom. The correct measure of the residuals, however, is

\[ e = y - X \hat{b} \]

where \( X \) is the matrix \( F \) with the columns comprised of the fitted values of endogenous regressors replaced by the actual values. This leads to a correct estimate of the sum of squared residuals, divided by the degrees of freedom, which we will simply call \( s^2 \). Since \( \hat{s}^2 \) enters the formula for the variance-covariance matrix of the coefficients, that matrix must be multiplied by \( s^2 / \hat{s}^2 \) to correctly estimate it. And the coefficient standard-errors produced by a second-stage OLS regression must be multiplied by \( s / \hat{s} \).

\(^3\)It turns out that the numerator is distributed as Chi-Square with \( n \) degrees of freedom and the denominator is distributed as Chi-Square with \( df \) degrees of freedom. An F-statistic is thus the ratio of two Chi-Square statistics.
To correctly produce second stage results we construct an XLispStat function called TSLS2 which takes as its five arguments:

1) a list representing the dependent variable.
2) a matrix of the independent variables with the fitted values from the first stage replacing the actual values of the endogenous variables in its first columns.
3) a matrix of independent variables identical to the one above except that the fitted values of the endogenous variables in the first columns are replaced by their actual values.
4) a matrix of the relevant instrumental variables.
5) an integer indicating the number of exogenous variables.

The code for the function is appropriately commented to explain what is being done, so it need not be explained in detail here. It contains a test of the relevance of the instruments based on the fact that if the instruments are uncorrelated with the error term their coefficients should be statistically insignificant in a regression of the second-stage residuals on those instruments and the other exogenous variables in the second stage-regression. This over-identifying-restrictions test involves constructing a standard F-Statistic of the restriction imposed by dropping the instruments from the regression and then multiplying that F-Statistic by the number of instruments to produce a J-Statistic which is distributed as Chi-Square with degrees of freedom equal to the number of instruments minus the number of endogenous variables. Accordingly, the test will only be performed when the number of instruments exceeds the number of exogenous variables. A statistically significant J-Statistic will indicate that the instruments are correlated with the residuals, casting doubt on their relevance. Finally, our TSLS2 function performs a Breusch-Pagan test for heteroskedasticity of, and a Breush-Godfrey test for serial correlation of, the residuals using functions developed and discussed in earlier Lessons. As in the case of all our regression functions, a text object called regressand, giving the name of the dependent variable, and a list of text objects called regressors, giving the names of the independent variables, must be in the work place when the function is called.

The code for performing the second stage regression for the supply curve in the example above is

```lisp
(def regressand "Quantity of Loans")
(def regressors (list "Constant" "Prime Rate" "T-Bill Rate" "Bank Deposits"))
(TSLS2 QLOANS (bind-columns FITPRATE TBRATE TOTBDEP)
(bind-columns PRIMRATE TBRATE TOTBDEP)(bind-columns CBNDRATE INDPROD) 1)
```

12
TWO-STAGE LEAST SQUARES: SECOND STAGE REGRESSION

Dependent Variable: Quantity of Loans

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-stat</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-88.023</td>
<td>13.966</td>
<td>-6.303</td>
<td>0.000</td>
</tr>
<tr>
<td>Prime Rate</td>
<td>6.900</td>
<td>1.901</td>
<td>3.629</td>
<td>0.000</td>
</tr>
<tr>
<td>T-Bill Rate</td>
<td>-7.080</td>
<td>2.272</td>
<td>-3.116</td>
<td>0.001</td>
</tr>
<tr>
<td>Bank Deposits</td>
<td>0.334</td>
<td>0.008</td>
<td>42.946</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Number of Observations: 72
Degrees of Freedom: 68
R-Squared: 0.9667502662640977
Sum of Squared Errors: 8633.59958956706
Residual Standard Error: 11.267861369945253
F-Statistic: 7426.005758901581
P-Value 0.0

Over-Identifying Restrictions Test

Chi-Square(1) Statistic = 4.577309420702708
P-Value = 0.032398047732226765

Breusch-Pagan ChiSquare Statistic: 11.93945948502644
P-Value = 0.007593450328177931

LM-Test for Serial Correlation of Residuals:
Number of Lags = 3
Chisquare Statistic = 33.80927665339088
P-Value = 2.173629508606112E-7
As you can see by comparison of this result with the standard OLS estimation of the supply equation on page 8, the signs of the coefficients do not change but their absolute magnitudes increase and their standard errors also increase. The prime-rate and t-bill-rate variables turn out to be more statistically significant in the two-stage case than when the supply relation is estimated in standard fashion. Notice also that the instruments fail the over-identifying restrictions test, casting serious doubt about their relevance and there is substantial heteroskedasticity and serial correlation in the residuals.4

The two-stage results in the case of the demand curve are as follows.

TSLS -- SECOND STAGE -- DEMAND CURVE USING THE TSLS2 FUNCTION
******************************************************************************

TWO-STAGE LEAST SQUARES: SECOND STAGE REGRESSION

Dependent Variable: Quantity of Loans

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-stat</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-204.332</td>
<td>74.196</td>
<td>-2.754</td>
</tr>
<tr>
<td>Prime Rate</td>
<td>-20.097</td>
<td>1.596</td>
<td>-12.589</td>
</tr>
<tr>
<td>Corp. Bond Rate</td>
<td>40.556</td>
<td>2.829</td>
<td>14.337</td>
</tr>
<tr>
<td>Ind. Prod.</td>
<td>2.306</td>
<td>0.453</td>
<td>5.095</td>
</tr>
</tbody>
</table>

Number of Observations: 72
Degrees of Freedom: 68
R-Squared: 0.748318008764265
Sum of Squared Errors: 65351.54698054201
Residual Standard Error: 31.00084130275433
F-Statistic: 2089.273008033983
P-Value 0.0

4These three test statistics were not calculated and reported by Maddala.
Over-Identifying Restrictions Test

Chi-Square(1) Statistic = 29.088166010841125
P-Value = 6.915813732799592E-8

Breusch-Pagan ChiSquare Statistic: = 8.241560943899907
P-Value = 0.04127446411754887

LM-Test for Serial Correlation of Residuals:
Number of Lags = 3
Chisquare Statistic = 112.48806269013818
P-Value = 0.0

As in the case of the supply curve the two-stage estimation, in comparison to the standard OLS result on page 7, increases the absolute magnitudes of all the coefficients without changing their signs. And again, the instruments fail the over-identifying restrictions test and there is heteroskedasticity and serial correlation in the residuals.

An obvious extension of our two-stage least squares analysis is to adjust for heteroskedasticity and serial correlation in the residuals by incorporating White and Newy-West procedures for making appropriate adjustments of the standard errors of the coefficients. To do this we extract from our OLS function the code that performs these adjustments and place it in a function of its own called HACSERR which takes as its two arguments the number of lags for dealing with serial correlation and the integer unity if the previous regression that is being analyzed is a two-stage one or a zero otherwise. If the number of lags is set at zero the function simply performs a White test to adjust for heteroskedasticity. If the regression being adjusted is the second stage regression from our TSLS2 function, our HACSERR function automatically further adjusts the standard errors upward by the magnitude of adjustment required by the TSLS2 function, using the ratio stderrat left in memory by that function. Also, unlike the HAC adjustment produced by our OLS function, the HACSERR function will not automatically select a lag of $0.75T^{1/3}$ based on its evaluation of the degree of serial correlation in the residuals when the lag argument is set at zero. The exact lag length must be imposed
when calling the function. Also, we can use our \texttt{HACERR} function to adjust the coefficient standard errors produced by our \texttt{OLSreg} function by simply calling it immediately thereafter. The additional code line and results from calling \texttt{HACERR} immediately after \texttt{TLS2} follow below for the above demand curve and supply curve estimates.

\texttt{(HACERR 3 1)}

**HAC MODIFICATIONS TO ABOVE RESULTS [FOR DEMAND CURVE]**

Coefficient Standard Errors are Newey-West HAC Adjusted with Lag = 3

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-stat</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-204.332</td>
<td>190.071</td>
<td>-1.075</td>
</tr>
<tr>
<td>Prime Rate</td>
<td>-20.097</td>
<td>4.260</td>
<td>-4.718</td>
</tr>
<tr>
<td>Corp. Bond Rate</td>
<td>40.556</td>
<td>6.608</td>
<td>6.137</td>
</tr>
<tr>
<td>Ind. Prod.</td>
<td>2.306</td>
<td>1.159</td>
<td>1.990</td>
</tr>
</tbody>
</table>

\texttt{(HACERR 3 1)}

**HAC MODIFICATIONS TO ABOVE RESULTS [FOR SUPPLY CURVE]**

Coefficient Standard Errors are Newey-West HAC Adjusted with Lag = 3

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-stat</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-88.023</td>
<td>25.351</td>
<td>-3.472</td>
</tr>
<tr>
<td>Prime Rate</td>
<td>6.900</td>
<td>3.122</td>
<td>2.210</td>
</tr>
<tr>
<td>T-Bill Rate</td>
<td>-7.080</td>
<td>4.128</td>
<td>-1.715</td>
</tr>
<tr>
<td>Bank Deposits</td>
<td>0.334</td>
<td>0.015</td>
<td>22.639</td>
</tr>
</tbody>
</table>

In both cases the coefficient standard errors are substantially increased and the corresponding t-ratios substantially reduced.

To run two-stage least squares estimates in the free program \texttt{R} we use the following two lines of code, both of which are in our script file \texttt{simlbias.R}. 

16
tsdemreg <- tsls(QLOANS~PRIMRATE+CBNDRATE+INDPROD, \
~CBNDRATE+INDPROD+TBRATE+TOTDEP)
summary(tsdemreg)

The code segment QLOANS PRIMRATE+CBNDRATE+INDPROD, represents the fundamental regression that is being run by the two-stage method and the code segment CBNDRATE+INDPROD+TBRATE+TOTDEP represents the exogenous variables and the instruments in that order. This tells R to run both stages of our demand curve regression. The code for the supply curve regression can be similarly constructed. R does not run an over-identifying restrictions test so we have to program one ourselves. The added code for this is as follows for the demand curve regression.

demresids <- tsdemreg$residuals 
#
OIRdemreg1 <- lm(demresids~CBNDRATE+INDPROD+TBRATE+TOTDEP)
summary(OIRdemreg1)
USSEdem <- deviance(OIRdemreg1)
OIRdemreg2 <- lm(demresids~CBNDRATE+INDPROD)
summary(OIRdemreg2)
RSSEdem <- deviance(OIRdemreg2)
Fdem <- ((RSSEdem-USSEdem)/2)/(USSEdem/OIRdemreg1$df)
Fdem
Jdem <- 2*Fdem
Jdem
PVdem <- 1 - pchisq(Jdem,df=1)
PVdem
The following is a selection of the results obtained in the output file `simulbias.Rou`.

2SLS Estimates

Model Formula: `QLOANS ~ PRIMRATE + CBNDRATE + INDPROD`

Instruments: `~CBNDRATE + INDPROD + TBRATE + TOTDEP`

Residuals:

<table>
<thead>
<tr>
<th></th>
<th>Min.</th>
<th>1st Qu.</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Qu.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residuals</td>
<td>-6.60e+01</td>
<td>-2.37e+01</td>
<td>-2.41e+00</td>
<td>1.54e-11</td>
<td>1.96e+01</td>
<td>7.30e+01</td>
</tr>
</tbody>
</table>

| Estimate      | Std. Error | t value | Pr(>|t|) |
|---------------|------------|---------|----------|
| (Intercept)   | -210.362   | 74.3037 | 6.095e-03|
| PRIMRATE      | -20.192    | 1.5997  | 0.000e+00|
| CBNDRATE      | 40.764     | 2.8402  | 0.000e+00|
| INDPROD       | 2.339      | 0.4526  | 2.246e-06|

Residual standard error: 30.9893 on 68 degrees of freedom

```r
> Fdem
[1] 14.30842
> Jdem <- 2*Fdem
> Jdem
[1] 28.61684
> PVdem <- 1 - pchisq(Jdem,df=1)
> PVdem
[1] 8.821147e-08
```

The R function above does not produce HAC adjusted coefficient standard errors. And we must be reluctant to use its `NeweyWest()` function to do so because we really cannot be sure how it will operate on the residuals from its `tsls()` function as compared to those from its standard OLS function `lm()`. Thankfully, the results above are very close to those produced in XLispStat by our `TSLS2` function.
Finally we perform the two-stage least squares analysis in \textit{Gretl}. This can be done by selecting \textit{Model} in the data window, then \textit{instrumental variables} and then \textit{Two-Stage Least Squares}, and following the instructions in the resulting window. Alternatively, we can simply construct the following \textit{Gretl} script file which we call \textit{simlbias.inp}.

```plaintext
# TWO-STAGE LEAST SQUARES ANALYSIS IN GRETL
#
open "E:\DSLMHTML\SIMLBIAS\loandata.gdt"  # modify accordingly
#
# DEMAND CURVE
#
ols QLOANS const PRIMRATE CBNDRATE INDPROD --robust
#
tsls QLOANS const PRIMRATE CBNDRATE INDPROD ; const CBNDRATE INDPROD TBRATE TOTDEP
tsls QLOANS const PRIMRATE CBNDRATE INDPROD ; const CBNDRATE INDPROD TBRATE TOTDEP --robust
#
# SUPPLY CURVE
#
ols QLOANS const PRIMRATE TBRATE TOTDEP --robust
#
tsls QLOANS const PRIMRATE TBRATE TOTDEP ; const TBRATE \ TOTDEP CBNDRATE INDPROD
tsls QLOANS const PRIMRATE TBRATE TOTDEP ; const TBRATE \ TOTDEP CBNDRATE INDPROD --robust
#
# FIRST-STAGE REGRESSION
#
ols PRIMRATE const TBRATE TOTDEP CBNDRATE INDPROD
```

Notice that the two-stage least squares function takes two sets of variables, each with a constant term, separated by the character ;. The first gives the basic regression we want to run and the second gives the set of instruments along with the independent variables. By adding the code segment --robust we can instruct \textit{Gretl} to adjust the coefficient standard errors for heteroskedasticity and serial correlation. The complete output, which is in the file \textit{simlbias.gou}, is shown below.
# TWO-STAGE LEAST SQUARES ANALYSIS IN GRETL

```plaintext
? open "E:\DSLMHTML\SIMLBIAS\loandata.gdt"  #modify appropriately

Read datafile E:\DSLMHTML\SIMLBIAS\loandata.gdt
observations range: 1979:01-1984:12

Listing 7 variables:
periodicity: 12, maxobs: 72
0) const 1) QLOANS 2) PRIMRATE 3) CBNDRATE 4) INDPROD
5) TBRATE 6) TOTDEP

# DEMAND CURVE
#

? ols QLOANS const PRIMRATE CBNDRATE INDPROD --robust

Model 1: OLS, using observations 1979:01-1984:12 (T = 72)
Dependent variable: QLOANS
HAC standard errors, bandwidth 3 (Bartlett kernel)

<table>
<thead>
<tr>
<th>coefficient</th>
<th>std. error</th>
<th>t-ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>-214.062</td>
<td>129.874</td>
<td>-1.648</td>
</tr>
<tr>
<td>PRIMRATE</td>
<td>-16.5616</td>
<td>1.86620</td>
<td>-8.875</td>
</tr>
<tr>
<td>CBNDRATE</td>
<td>36.6761</td>
<td>3.83334</td>
<td>9.568</td>
</tr>
<tr>
<td>INDPROD</td>
<td>2.36876</td>
<td>0.786843</td>
<td>3.010</td>
</tr>
</tbody>
</table>

Mean dependent var 361.5556  S.D. dependent var 62.31587
Sum squared resid 63305.98  S.E. of regression 30.51181
R-squared 0.770391  Adjusted R-squared 0.760261
F(3, 68) 45.27016  P-value(F) 3.38e-16
Log-likelihood -346.2101  Akaike criterion 700.4201
Schwarz criterion 709.5268  Hannan-Quinn 704.0455
rho 0.664409  Durbin-Watson 0.658519

20
Model 2: TSLS, using observations 1979:01-1984:12 (T = 72)
Dependent variable: QLOANS
Instrumented: PRIMRATE
Instruments: const CBNDRATE INDPROD TBRATE TOTDEP

<table>
<thead>
<tr>
<th></th>
<th>coefficient</th>
<th>std. error</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>-220.874</td>
<td>77.9247</td>
<td>-2.834</td>
<td>0.0046  ***</td>
</tr>
<tr>
<td>PRIMRATE</td>
<td>-20.8050</td>
<td>1.67702</td>
<td>-12.41</td>
<td>2.43e-035 ***</td>
</tr>
<tr>
<td>CBNDRATE</td>
<td>41.4223</td>
<td>2.97754</td>
<td>13.91</td>
<td>5.39e-044 ***</td>
</tr>
<tr>
<td>INDPROD</td>
<td>2.42085</td>
<td>0.474417</td>
<td>5.103</td>
<td>3.35e-07 ***</td>
</tr>
</tbody>
</table>

Mean dependent var 361.5556
S.D. dependent var 62.31587
Sum squared resid 71759.49
S.E. of regression 32.48517
R-squared 0.761451
Adjusted R-squared 0.750927
F(3, 68) 77.72069
P-value(F) 6.24e-22
rho 0.626987
Durbin-Watson 0.735020

Hausman test -
Null hypothesis: OLS estimates are consistent
Asymptotic test statistic: Chi-square(1) = 81.7343
with p-value = 1.55665e-019

Sargan over-identification test -
Null hypothesis: all instruments are valid
Test statistic: LM = 19.6841
with p-value = P(Chi-Square(1) > 19.6841) = 9.13602e-006

Weak instrument test -
First-stage F-statistic (2, 67) = 133.379
Critical values for desired TSLS maximal size, when running
tests at a nominal 5% significance level:
size 10% 15% 20% 25%
value 19.93 11.59 8.75 7.25
Maximal size is probably less than 10%
Model 3: TSLS, using observations 1979:01-1984:12 (T = 72)
Dependent variable: QLOANS
Instrumented: PRIMRATE
Instruments: const CBNDRATE INDPROD TBRATE TOTDEP
HAC standard errors, bandwidth 3 (Bartlett kernel)

<table>
<thead>
<tr>
<th></th>
<th>coefficient</th>
<th>std. error</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>-220.874</td>
<td>111.602</td>
<td>-1.979</td>
<td>0.0478  **</td>
</tr>
<tr>
<td>PRIMRATE</td>
<td>-20.8050</td>
<td>2.28982</td>
<td>-9.086</td>
<td>1.03e-019 ***</td>
</tr>
<tr>
<td>CBNDRATE</td>
<td>41.4223</td>
<td>3.80076</td>
<td>10.90</td>
<td>1.17e-027 ***</td>
</tr>
<tr>
<td>INDPROD</td>
<td>2.42085</td>
<td>0.674790</td>
<td>3.588</td>
<td>0.0003  ***</td>
</tr>
</tbody>
</table>

Mean dependent var 361.5556
S.D. dependent var 62.31587
Sum squared resid 71759.49
S.E. of regression 32.48517
R-squared 0.761451
Adjusted R-squared 0.750927
F(3, 68) 51.46455
P-value(F) 1.78e-17
rho 0.626987
Durbin-Watson 0.735020

Hausman test -
Null hypothesis: OLS estimates are consistent
Asymptotic test statistic: Chi-square(1) = 81.7343
with p-value = 1.55665e-019

Sargan over-identification test -
Null hypothesis: all instruments are valid
Test statistic: LM = 19.6841
with p-value = P(Chi-Square(1) > 19.6841) = 9.13602e-006

Weak instrument test -
First-stage F-statistic (2, 67) = 89.2323
A value < 10 may indicate weak instruments
The Hausman test that Gretl runs in its `tsls` function tests the null hypothesis that the coefficients in a simple OLS regression are consistent. Failure to reject that null hypothesis means that two-stage-least squares analysis is unnecessary. The interpretation of the other two tests should be obvious.

```
# SUPPLY CURVE
#
? ols QLOANS const PRIMRATE TBRATE TOTDEP --robust
```

Model 4: OLS, using observations 1979:01-1984:12 (T = 72)
Dependent variable: QLOANS
HAC standard errors, bandwidth 3 (Bartlett kernel)

<table>
<thead>
<tr>
<th>coefficient</th>
<th>std. error</th>
<th>t-ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>-78.9066</td>
<td>13.1195</td>
<td>-6.014</td>
</tr>
<tr>
<td>PRIMRATE</td>
<td>2.05005</td>
<td>1.17351</td>
<td>1.747</td>
</tr>
<tr>
<td>TBRATE</td>
<td>-1.78460</td>
<td>1.50184</td>
<td>-1.188</td>
</tr>
<tr>
<td>TOTDEP</td>
<td>0.336810</td>
<td>0.00766580</td>
<td>43.94</td>
</tr>
</tbody>
</table>

Mean dependent var 361.5556 S.D. dependent var 62.31587
Sum squared resid 11833.13 S.E. of regression 13.19154
R-squared 0.957082 Adjusted R-squared 0.955188
F(3, 68) 656.8475 P-value(F) 4.02e-50
Log-likelihood -285.8353 Akaike criterion 579.6706
Schwarz criterion 588.7773 Hannan-Quinn 583.2960
rho 0.135962 Durbin-Watson 1.716490

Excluding the constant, p-value was highest for variable 5 (TBRATE)
Model 5: TSLS, using observations 1979:01-1984:12 (T = 72)
Dependent variable: QLOANS
Instrumented: PRIMRATE
Instruments: const TBRATE TOTDEP CBNDRATE INDPROD

<table>
<thead>
<tr>
<th></th>
<th>coefficient</th>
<th>std. error</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>-89.6899</td>
<td>18.1217</td>
<td>-4.949</td>
<td>7.45e-07 ***</td>
</tr>
<tr>
<td>PRIMRATE</td>
<td>6.63045</td>
<td>2.46133</td>
<td>2.694</td>
<td>0.0071 ***</td>
</tr>
<tr>
<td>TBRATE</td>
<td>-7.08268</td>
<td>2.94262</td>
<td>-2.407</td>
<td>0.0161 **</td>
</tr>
<tr>
<td>TOTDEP</td>
<td>0.339387</td>
<td>0.0100972</td>
<td>33.61</td>
<td>1.13e-247 ***</td>
</tr>
</tbody>
</table>

Mean dependent var 361.5556 S.D. dependent var 62.31587
Sum squared resid 14551.46 S.E. of regression 14.62847
R-squared 0.947431 Adjusted R-squared 0.945112
F(3, 68) 412.6124 P-value(F) 1.50e-43
rho 0.209809 Durbin-Watson 1.575911

Hausman test -
Null hypothesis: OLS estimates are consistent
Asymptotic test statistic: Chi-square(1) = 6.78328
with p-value = 0.00920157

Sargan over-identification test -
Null hypothesis: all instruments are valid
Test statistic: LM = 3.87846
with p-value = P(Chi-Square(1) > 3.87846) = 0.0489094

Weak instrument test -
First-stage F-statistic (2, 67) = 12.5559
Critical values for desired TSLS maximal size, when running tests at a nominal 5% significance level:

<table>
<thead>
<tr>
<th>size</th>
<th>10%</th>
<th>15%</th>
<th>20%</th>
<th>25%</th>
</tr>
</thead>
<tbody>
<tr>
<td>value</td>
<td>19.93</td>
<td>11.59</td>
<td>8.75</td>
<td>7.25</td>
</tr>
</tbody>
</table>
Maximal size may exceed 10%
Model 6: TSLS, using observations 1979:01-1984:12 (T = 72)
Dependent variable: QLOANS
Instrumented: PRIMRATE
Instruments: const TBRATE TOTDEP CBNDRATE INDPROD
HAC standard errors, bandwidth 3 (Bartlett kernel)

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>Z</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>-89.6899</td>
<td>18.7285</td>
<td>-4.789</td>
</tr>
<tr>
<td>PRIMRATE</td>
<td>6.63045</td>
<td>2.36423</td>
<td>2.804</td>
</tr>
<tr>
<td>TBRATE</td>
<td>-7.08268</td>
<td>3.28435</td>
<td>-2.156</td>
</tr>
<tr>
<td>TOTDEP</td>
<td>0.339387</td>
<td>0.0104725</td>
<td>32.41</td>
</tr>
</tbody>
</table>

Mean dependent var 361.5556 S.D. dependent var 62.31587
Sum squared resid 14551.46 S.E. of regression 14.62847
R-squared 0.947431 Adjusted R-squared 0.945112
F(3, 68) 374.2501 P-value(F) 3.45e-42
rho 0.209809 Durbin-Watson 1.575911

Hausman test -
Null hypothesis: OLS estimates are consistent
Asymptotic test statistic: Chi-square(1) = 6.78328
with p-value = 0.00920157

Sargan over-identification test -
Null hypothesis: all instruments are valid
Test statistic: LM = 3.87846
with p-value = P(Chi-Square(1) > 3.87846) = 0.0489094

Weak instrument test -
First-stage F-statistic (2, 67) = 8.08769
A value < 10 may indicate weak instruments
# FIRST-STAGE REGRESSION

Model 7: OLS, using observations 1979:01-1984:12 (T = 72)
Dependent variable: PRIMRATE

<table>
<thead>
<tr>
<th></th>
<th>coefficient</th>
<th>std. error</th>
<th>t-ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>1.95945</td>
<td>3.03128</td>
<td>0.6464</td>
<td>0.5202</td>
</tr>
<tr>
<td>TBRATE</td>
<td>0.772759</td>
<td>0.116738</td>
<td>6.620</td>
<td>7.22e-09 ***</td>
</tr>
<tr>
<td>TOTDEP</td>
<td>-0.00521408</td>
<td>0.00151897</td>
<td>-3.433</td>
<td>0.0010 ***</td>
</tr>
<tr>
<td>CBNDRATE</td>
<td>0.760620</td>
<td>0.178192</td>
<td>4.269</td>
<td>6.34e-05 ***</td>
</tr>
<tr>
<td>INDPDO</td>
<td>0.00705941</td>
<td>0.0224958</td>
<td>0.3138</td>
<td>0.7546</td>
</tr>
</tbody>
</table>

Mean dependent var | 14.08333  S.D. dependent var | 3.176402
Sum squared resid  | 94.24382  S.E. of regression | 1.186012
R-squared           | 0.868440  Adjusted R-squared   | 0.860586
F(4, 67)            | 110.5684  P-value(F)            | 9.31e-29
Log-likelihood      | -111.8555 Akaike criterion     | 233.7109
Schwarz criterion   | 245.0943  Hannan-Quinn          | 238.2427
rho                 | 0.650325  Durbin-Watson         | 0.701157

Excluding the constant, p-value was highest for variable 4 (INDPROD)

When the first stage regression is run in Gretl it confirms the results from XLispStat and R that the industrial production variable is a weak instrument. With respect to the two-stage regression results, where no adjustment is made for heteroskedasticity and autocorrelation, Gretl gives a similar result to those using XLispStat and R in estimating the demand curve, but its coefficient standard errors are substantially higher, and t-ratios substantially lower in its estimation of the supply curve. When robust estimation is specified, Gretl produces much lower coefficient standard errors and higher t-ratios than were obtained using our HACSEERR function in XLispStat in both the demand and supply curve estimation. In fact, if we instruct our HACSEERR function to act as though it were making adjustments to the coefficient stan-
dard errors for a standard OLS function instead of the TSLSS2 function, we obtain standard errors and t-statistics quite similar to those produced by Gretl with robust estimation. Given this fact, it may be wise to always also produce HAC standard-error adjustments in the two-stage case using a zero rather than unity as the second argument to our HACERR function.

Of course, not all simultaneity bias problems involve joint estimation of supply and demand curves. We might want to determine whether a particular medical treatment prolongs lives and approach the problem by running a regression of peoples’ length of life on whether they have had the treatment, including additional exogenous variables to control for other factors that affect mortality such as lifestyle, age, weight and other health conditions. A simultaneity problem arises here because the people who select the treatment may decide based on advice of a family doctor and on other medical and other factors not in the data set. People who are more (or less) likely to die soon may be the ones that select the treatment, leading to bias in the statistical result. What is necessary is an instrument that is correlated with the treatment but uncorrelated with the omitted factors that affect survival. A

**Real Exchange Rates and Monetary Policy: Canada**

We now turn to an examination of the simultaneity bias problems arising in an international macroeconomic investigation of the determinants of real and nominal exchange rates and the consequences of government efforts to manipulate them. The real exchange rate, which is the relative price of domestic output in terms of foreign output, can be expressed as

\[ Q = \frac{\Pi P}{\tilde{P}} \]  

where \( Q \) is the real exchange rate, \( \Pi \) the nominal exchange rate defined as the foreign currency price of domestic currency, \( P \) the domestic price level and \( \tilde{P} \) the foreign price level. Given this real exchange rate, the nominal

---

5See pages 360 through 165 of the Stock and Watson book for a number of examples, one of which is an actual study of cardiac catheterization on which our discussion here is based.

exchange rate will be inversely related to the ratio of the domestic over the foreign price level—that is, by the extent of past domestic relative to foreign price inflation. The domestic and foreign price levels can be expressed as geometrically weighted averages of the prices of the traded and non-traded components of the respective countries’ outputs.

\[ P = P_N^\theta P_T^{1-\theta} \]  

and

\[ \tilde{P} = \tilde{P}_N^\tilde{\theta} \tilde{P}_T^{1-\tilde{\theta}} \]

where \(1 > \theta > 0\) and \(1 > \tilde{\theta} > 0\) are the fractions of domestic and foreign output represented by non-traded components. Here it is assumed that all goods have traded and non-traded components. 7 Substituting equations (7) and (8) into (6), we obtain

\[
Q = \frac{\Pi P_N^\theta P_T^{1-\theta}}{\tilde{P}_N^\theta \tilde{P}_T^{1-\theta}} = \frac{\Pi P_N^\theta (\tilde{P}_T D / \Pi)^{1-\theta}}{\tilde{P}_N^\theta \tilde{P}_T^{1-\theta}} = \frac{\Pi (\Pi P_N^\theta)^{1-\theta}}{\tilde{P}_N^\theta \tilde{P}_T^{1-\theta}} \left[ \frac{\tilde{P}_T D}{\tilde{P}_T^{1-\theta}} \right].
\]

(9)

The real exchange rate of Canada with respect to the United States will thus depend on the ratio of the prices of the non-traded components of Canadian output to the prices of the non-traded components of U.S. output and on the price of Canadian traded output components divided by the price of U.S. traded output components.

It is clear from the above that we can expect Canada’s real exchange rate with respect to the U.S. to rise when the prices of commodities and energy rise in international markets, relative to the prices of other goods, because production of these commodities represents a higher proportion of Canadian output than United States output. More broadly, we would expect that a rise in Canada’s terms of trade with respect to the rest of the world relative the U.S terms of trade with the rest of the world would also lead to an increase in the real exchange rate. And, according to the Balassa-Samuelson hypothesis,

7 Barbers may be using clippers imported from abroad, and exports of wheat (as well as imports of clippers) will have cost components representing domestic labour required to arrange transport and sale.
we would also expect the real exchange rate to rise in response to an increase in domestic relative to foreign full-employment income. As income rises so do real wages and the relative increase in real wages increases the cost of producing the non-traded—that is, labour intensive—components of output relative to the cost of producing foreign non-traded output components. A further obvious factor causing the real exchange rate to rise will be shifts of demand of domestic residents from goods with low non-traded components to those with high non-traded components. While shifts of this sort will be extremely difficult to measure, one obvious measurable factor might be the share of government expenditure in domestic output since there are obvious political pressures on government to channel its spending as directly as possible to domestic residents.

Finally, we can expect that a decision of international investors, in response to new technological developments, to increase their investment in Canada relative to their investment in United States will produce an increased demand for the non-traded components of Canadian output, requiring a higher relative price of that output to achieve equilibrium. This rise in the real exchange rate will have to increase the current account deficit sufficiently to offset the increased net capital inflow. This result follows from the fact that domestic output can be divided into the following components.

\[ Y = C + I + B_T + DSB \]  

(10)

where \( Y \) is income, \( C \) is total private plus government expenditure on consumption, \( I \) is total private plus government expenditure on investment, \( B_T \) is the balance of trade in goods and services excluding the services of capital, and \( DSB \) is the debt service balance which equals total income from foreign employed capital owned by domestic residents minus total income from domestically employed capital owned by foreigners. Subtraction of total consumption and investment from both sides produces the expression

\[ Y - C - I = B_T + DSB \]  

(11)

which reduces to

\[ S - I = CAB \]  

(12)

---

where $S = Y - C$ is the level of savings and $CAB = B_T + DSB$ is the current account balance. These conditions are true by definition when the variables are the actual values and represent the condition of output-market equilibrium—that is, the equality of aggregate demand and supply—when the variables are the desired magnitudes. For equilibrium to occur, the real exchange rate, and perhaps also the level of income and thereby savings must adjust to ensure that the above equality holds. The role of real exchange rate adjustment becomes obvious when we recognize that the current account balance can be expressed

$$CAB = B_T(Q, Y, \tilde{Y}) + DSB$$

(13)

where $\tilde{Y}$ is the level of foreign income. This expression can be written equivalently as

$$S - I - DSB = B_T(Q, Y, \tilde{Y})$$

(14)

or as

$$I - S + DSB = -B_T(Q, Y, \tilde{Y})$$

(15)

which states simply that the net capital inflow plus debt service balance must be equal to the negative of the balance of trade in goods and services. When capital flows in, a rise in the real exchange rate will be required to increase imports relative to exports and thereby decrease the balance of trade surplus or increase the balance of trade deficit to create a flow of goods into the country equal to the inflow flow of ownership claims to capital. An increase in domestic income will increase imports at any given real exchange rate, reducing the balance of trade, and an increase in foreign income will raise exports, increasing it. This will, of course, require corresponding adjustments of savings relative to investment—hence, the general equilibrium nature of the adjustment process and the resulting possibility of simultaneity bias in our econometric analysis.

The above analysis suggests that we begin by regressing the logarithm of Canada’s real exchange rate with respect to the United States $LREXCAUS$, calculated on the basis of equation (6) on the following variables:
LRPCOMXEN — the logarithm of commodity prices in U.S. dollars divided by an equally weighted average of the U.S. dollar prices of U.S. exports and imports.

LRPENERGY — the logarithm of energy prices in U.S. dollars divided by an equally weighted average of the U.S. dollar prices of U.S. exports and imports.

LRTOTCAUS — the logarithm of Canada’s terms of trade with respect to the rest of the world divided by the U.S. terms of trade with respect to the rest of the world.

DGCONCAUS — Canadian government consumption expenditure as a percentage of GDP minus U.S. government consumption expenditure as a percentage of GDP.

LRGDPCA — logarithm of Canadian real GDP.

LRGDPUS — logarithm of United States GDP.

DNCICAUS — the net capital inflow into Canada plus the debt service balance as a percentage of Canadian GDP minus the net capital inflow into the U.S. plus that country’s debt service balance as a percentage of U.S. GDP, where the net capital flows plus the debt service balances are estimated as the negative of the country’s balances of trade in goods and services. We should keep in mind here that the debt service balances will be determined by previous savings and investment levels and be therefore unresponsive to changes in current levels of the real exchange rate.

Our regression analysis begins using quarterly data for the period 1974 through 2010. It turns out that the only statistically significant variables are the logarithms of the real prices of commodities less energy and of energy and the excess of the net capital inflow into Canada as a percentage of GDP over the net capital inflow into the U.S. as a percentage of that country’s GDP. This regression and the statistical analysis that follows is done in the XLispStat batch file rexcaus.lsp and the output is in the file rexcaus.lou. Similar input and output files produced using R are rexcaus.R and rexcaus.Rou. In the case of Gretl we perform the calculations and obtain the results in a session file called rexcaus.gretl. The XLispStat output is as follows (and the outputs of the other two programs are similar.\(^9\)

\(^9\)We state the time-period of the regression in the output file by simply attaching additional code beyond the description of the dependent variable in the regressand object.
ORDINARY LEAST SQUARES REGRESSION

Dependent Variable: LREXCAUS --- Time Period 1974Q1 -- 2010Q4

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-stat</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>2.108</td>
<td>0.252</td>
<td>8.375</td>
<td>0.000</td>
</tr>
<tr>
<td>LRPCOMXEN</td>
<td>0.311</td>
<td>0.045</td>
<td>6.949</td>
<td>0.000</td>
</tr>
<tr>
<td>LRPENERGY</td>
<td>0.138</td>
<td>0.016</td>
<td>8.786</td>
<td>0.000</td>
</tr>
<tr>
<td>DNCICAUS</td>
<td>0.027</td>
<td>0.002</td>
<td>13.747</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Number of Observations: 148
Degrees of Freedom: 144
R-Squared: 0.7206317588499688
Adjusted R-Squared: 0.714815871593431
Sum of Squared Errors: 0.6039005016567894
F-Statistic: 123.81623724445541
P-Value 0.0

The failure of the real income variables to be statistically significant is not surprising given the very similar growth patterns through time in the two countries—there was not enough difference between their growth patterns to significantly affect the relative price of Canadian output in terms of output in the United States. It turns out, however, that our interest here is also on the question as to whether short-run business cycle shocks had any effect on real exchange rates. One way of investigating this is to include the Canadian and U.S. employment rates, EMPRCA and EMPRUS, as variables. These are defined simply as 100 minus the percentages of the labour forces unemployed. When we include these along with the logarithms of the two countries’ real GDPs we obtain the surprising result below. Because of the unavailability of appropriate employment rate series prior to 1976, it was necessary to shorten the time-period to begin at that year. Also, we use our OLS function to incorporate HAC coefficient standard errors.10

10It also turns out that the real GDP variables must be obtained by deflating the nominal series by the implicit GDP deflators rather than by the consumer price indexes to obtain statistically significant results. This would seem appropriate on the grounds that the GDP deflators should be a better measure of the price of output.
### ORDINARY LEAST SQUARES REGRESSION

Dependent Variable: LREXCAUS  ---  Time Period 1976Q1 -- 2010Q4

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-stat</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.209</td>
<td>0.647</td>
<td>-0.322</td>
</tr>
<tr>
<td>LRPCOMXEN</td>
<td>0.583</td>
<td>0.103</td>
<td>5.674</td>
</tr>
<tr>
<td>LRPENERGY</td>
<td>0.188</td>
<td>0.034</td>
<td>5.480</td>
</tr>
<tr>
<td>DNCICAUS</td>
<td>0.032</td>
<td>0.005</td>
<td>6.206</td>
</tr>
<tr>
<td>LRGDPPCA</td>
<td>1.991</td>
<td>0.469</td>
<td>4.248</td>
</tr>
<tr>
<td>LRGDPUS</td>
<td>-1.702</td>
<td>0.409</td>
<td>-4.159</td>
</tr>
<tr>
<td>EMPRCA</td>
<td>-0.038</td>
<td>0.009</td>
<td>-4.081</td>
</tr>
<tr>
<td>EMPRUS</td>
<td>0.029</td>
<td>0.010</td>
<td>2.744</td>
</tr>
</tbody>
</table>

Number of Observations: 140
Degrees of Freedom: 132
R-Squared: 0.8123470578892383
Adjusted R-Squared: 0.8023957655045767
Sum of Squared Errors: 0.36693459973032794
F-Statistic: 81.63231734014238
P-Value 0.0

Coefficient Standard Errors are Newey-West HAC Adjusted with Lag = 3

Breusch-Pagan ChiSquare Statistic: = 41.51027532002726
  P-Value = 6.459953220128156E-7

LM-Test for Serial Correlation of Residuals:
  Number of Lags = 3
  Chisquare Statistic = 216.15917351030717
  P-Value = 0.0
The coefficients of real commodity and energy prices and the real net capital inflow variable are a bit larger when the four new variables are added and the results are now consistent with the Balassa-Samuelson hypothesis. An increase in Canadian relative to United States full-employment income levels, leads to an increase in the relative cost and price of the non-traded components of Canadian output relative to U.S. output causing the real exchange rate to rise. At the same time, given these long-run output levels, a short-term increase in Canadian employment relative to the U.S. reduces the relative price of Canadian output in terms of U.S. output by increasing its relative supply.

Figure 5: The effect on the exchange rate of the excess of the price of energy in U.S. dollars relative to the average U.S. dollar price of U.S. exports and imports.
The figure above shows the effect of energy prices on the real exchange rate with the average level of that effect adjusted to conform to the average level of the real exchange rate. The importance of higher energy prices in generating upward movement of the real exchange rate in the period after 2000 is quite clear. The pattern of effects of the commodity price variable excluding energy on the real exchange rate is not easily seen from a graph and is therefore not shown although that effect is unquestionably statistically significant in our regression.

**Figure 6:** The effect on the exchange rate of the excess of the Canadian net capital inflow as a percentage of domestic GDP over the U.S. net capital inflow as a percentage of GDP.

The effect on the real exchange rate of the difference of the net capital inflow relative to GDP in Canada over the net capital inflow relative to GDP in the United States is shown in Figure 6 above. The reduction in the net
capital inflow, or increase in the net outflow, was obviously an important factor in the more than 20 percent decline of the real exchange rate between 1974 and the mid-to-late 1980s and the roughly equivalent increase by the early 1990s and then the subsequent decline of around 40 percent by around 2001 followed by a return to around the early-1990s level by 2010. Indeed it was surely the major factor determining movements of the real exchange rate of Canada with respect to the United States during the period under examination. But can we be sure that the causation runs from capital flow changes to real exchange rate changes and not the reverse? Or could it be running simultaneously in both directions? To investigate these issues we need to more carefully develop the underlying economic theory.

Figure 7: Real Exchange Rate Determination

Figure 7 above presents an economic analysis of the underlying causal relationships in real exchange rate determination. Think in terms of the long-run with full employment. Clearly, an increase in world investment in Canadian employed capital with unchanged ratios of domestic and foreign savings to the respective GDPs will cause the real exchange rate of Canada with respect to the United States to rise. The increased purchases of Canadian assets by world residents and the associated increase in domestic investment relative to savings will involve increased purchases of Canadian
dollars in the foreign exchange market. The Canadian dollar will appreciate as the equilibrium moves to the right along the negatively sloped BT curve from point a to point c. This will increase the level of imports relative to exports by an amount equal to the increase in domestic investment relative to savings. Also, an increase in the world demand for energy and other commodities will cause an increase in world prices of these goods. Since these goods represent a higher fraction of output in Canada than elsewhere, the world price of Canadian output will rise relative to aggregate output prices in the United States and elsewhere. At a given nominal exchange rate, this will imply a rise in the Canadian relative to the foreign price level. To the extent that the Canadian authorities maintain a monetary policy that prevents an increase in the domestic price level and domestic-currency prices of oil, natural gas and other commodities, the resulting excess world demand for Canadian output will represent an excess demand for the Canadian dollar, causing the dollar to appreciate relative to other currencies including the U.S. dollar. In either case the effect will be a shift of the BT curve to the right and a rise in Canada’s real exchange rate with respect to the U.S. (and other countries) as the equilibrium moves from point a to point b. Also, under full employment conditions the BT curve will shift to the right if Canadian output and income rise relative to output and income abroad as specified by the Balassa-Samuelson hypothesis.

Now suppose that short-run changes occur in the level of employment in Canada relative to the U.S. The resulting temporary increase in the supply of Canadian relative to foreign output will cause the price of Canadian output to fall relative to the price of U.S. and rest-of-world output, shifting the BT curve downward to the left.

If the SI line is vertical and exogenous shifts in net capital flows are substantial, and our least-squares regression includes all variables that substantially affect the real exchange rate, our regression will give us an unbiased estimate of the slope of the BT line. Given additional factors other than energy and commodity prices that affect the demand for Canadian exports relative to imports, the scatter plot of equilibrium points will be the area a b c d in Figure 8 below when the SI line and BT curve shift within the ranges indicated. It is clear that minimizing the squared vertical deviations of the data points from the regression line will produce a line lying on top of the BT line. But is it reasonable to assume that the SI line is vertical? To the extent that it is positively or negatively sloped, the slope of the estimated
BT line will be flatter or steeper than the true line because of simultaneity bias.

It would seem reasonable to expect that a change in the real exchange rate would have effects on the net capital inflow in both directions. The fact that the relative price of Canadian in terms of world output increases with a rise in the real exchange rate can be reasonably expected to make additional investment in Canada profitable. At the same time, however, the effect of the rise in the real exchange rate on the wealth of owners of Canadian holders of domestic capital will probably result in some increase in domestic savings. The fact that Canadian investment and savings move in the same direction makes the combined effect on the slope of the SI curve unclear. Given our lack of knowledge about the magnitudes of the opposing effects, the best alternative would seem to be to continue assuming that the SI line is vertical, recognizing that some bias in our results in one direction or the other may be present.

At this point it is important to recall that our net capital inflow variable is the net capital inflow into Canada from all countries as a percentage of Canadian GDP minus the net capital inflow from all countries into the U.S.
a percentage of that country’s GDP. And these net capital flow variables must equal imports from minus exports to all countries. Since the net capital inflows into both Canada and the United States can arise as a consequence of technological and other developments in third countries, and capital flows into and out of those countries, a serious additional source of simultaneity bias in our coefficient estimates in the regressions above is present. The correlation of capital flows into Canada with that country’s real exchange rate with respect to the United States will depend on what is happening in third countries to influence their capital inflows and outflows. The real exchange rates of those other countries with respect to the United States will also be changing through time so that a change in Canada’s real exchange rate with respect to the United States will be correlated with changes in the real exchange rates of other countries with respect to the United States and will therefore not necessarily be a good measure of Canada’s real exchange rate with respect to the rest of the world. Analysis based on the regression coefficient of the net capital inflow variable cannot routinely hold constant the real exchange rates of the U.S. with respect to third countries to specify that changes in Canada’s real exchange rate with respect to the United States represent proportional changes in her real exchange rate with respect to the rest of the world. Our Canadian import, export and net capital flow variables are with respect to the entire world outside Canada, not just the United States.

To circumvent this problem we need to include in our regression the real exchange rates with respect to the United States of all the major trading countries in the world. This is, of course, impossible because of resulting degrees-of-freedom limitations. It turns out, however, that when the real exchange rates of the United Kingdom and Japan with respect to the United States are added to our regression individually, their coefficients turn out to be statistically insignificant and, when included together, both coefficients are statistically insignificant. We have insufficient data to add the real exchange rate of the Euro Area with respect to the United States over the entire period. For the period 1999 onward, however, we can usefully add the real exchange rates of Japan, the Euro Area and the United Kingdom with respect to the U.S. to a regression that does not include real GDPs and employment levels. When we do this the real exchange rates of the Euro Area and the United Kingdom, but not Japan, turn out to be statistically significant when added separately but the Euro Area and U.K. real exchange
rates both are statistically insignificant when added together. While the sample size is small, bootstrapping to obtain coefficient estimates produces a range of coefficient values consistent with that obtained by standard OLS. The XLispStat code for the regression analysis that includes the Euro Area and U.K. real exchange rates is

```
(def regressand "LREXCAUS --- Time Period 1999Q1 -- 2010Q4")
(def regressors (list "Constant" "LRPCOMXEN" "LRPENERGY" "DNCICAUS" "LREXUKUS" "LREXEUUS"))
(OLSreg (bind-columns (remove-first 108 lrexcaus))(bind-columns (remove-first 108 lrpcomxen)(remove-first 108 lrpenergy) (remove-first 108 dncicaus)(remove-first 100 lrexukus) (log rexeuus)) 1)
(def USSE OLSSSE)
(def udf OLSdf)
(OLSBS 1000 (list " Constant" " LRPCOMXEN" " LRPENERGY" " DNCICAUS" " LREXUKUS" " LREXEUUS")

where the residual sum of squares and the degree of freedom are saved for future use. The results are as follows.

ORDINARY LEAST SQUARES REGRESSION

Dependent Variable: LREXCAUS --- Time Period 1999Q1 -- 2010Q4

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-stat</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.403</td>
<td>0.304</td>
<td>-1.326</td>
</tr>
<tr>
<td>LRPCOMXEN</td>
<td>0.465</td>
<td>0.085</td>
<td>5.464</td>
</tr>
<tr>
<td>LRPENERGY</td>
<td>0.227</td>
<td>0.024</td>
<td>9.512</td>
</tr>
<tr>
<td>DNCICAUS</td>
<td>0.022</td>
<td>0.004</td>
<td>5.139</td>
</tr>
<tr>
<td>LREXUKUS</td>
<td>0.184</td>
<td>0.145</td>
<td>1.268</td>
</tr>
<tr>
<td>LREXEUUS</td>
<td>0.023</td>
<td>0.126</td>
<td>0.184</td>
</tr>
</tbody>
</table>

Number of Observations: 48
Degrees of Freedom: 42
R-Squared: 0.9672678469701822
Adjusted R-Squared: 0.96337116208568
The bootstrapped coefficient values of the three real exchange rate variables bracket zero and the coefficient of the real net capital inflow variable ranges from 0.0138 to 0.0308, bracketing the values in the above regression. The values of this coefficient obtained in the full-period regressions, however, turn out to be roughly equivalent to the maximum value obtained in the above bootstrapped regression. The regression coefficient of the net capital inflow variable in the above regression is about two-thirds the size of the coefficients obtained in the regressions for the period 1976Q1 to 2010Q4. It was feasible to include the real GDPs and unemployment rates in the longer-period regressions and, it turns out, the omission of those variables in from those regressions reduces the coefficient of the net capital inflow variable to roughly the levels obtained in the comparable short-period regressions. It is reasonable to conclude, therefore, that the more appropriate measure of the net capital flow coefficient is the 0.032 obtained in the longer period regression that contained the Real GDP and unemployment rate variables but no third-country real exchange rate variables.
The plot of the Euro Area and United Kingdom real exchange rates with respect to the United States, shown in Figure 9 above, indicates that they tend to move in step. This suggests the possibility that the two exchange rates, when added as independent variables, are so highly correlated that we have a problem of multicollinearity that might be handled by including one of the real exchange rates as the representative of both. To check for this we perform F-tests of the restrictions imposed by removing the Euro Area and U.K. real exchange rates from the 1999Q1-2010Q4 regression that did not include the real GDP and employment rate variables. The code for doing this is as follows.

Figure 9: The real exchange rates of the Euro Area and the United Kingdom with respect to the United States, both indexed on a base of 2000 = 100.
(def regressand "LREXCAUS  ---  Time Period 1999Q1 -- 2010Q4")
(def regressors (list "Constant" "LRPCOMXEN" "LRPENERGY" "DNCICAUS"))
(OLSreg (bind-columns (remove-first 108 lrexcaus))
(bind-columns (remove-first 108 lrpcomxen)(remove-first 108 lrpenergy)
(remove-first 108 dncicaus)) 1)
(def RSSE OLSSSE)
(def FStat (/ (/ (- RSSE USSE) OLSdf)(/ USSE udf)))
(def PV (- 1 (f-cdf FSTAT 3 udf)))
(OLSBS 1000 (list " Constant" " LRPCOMXEN" " LRPENERGY"
"  DNCICAUS")
(terpri)
(princ "F-Test for exclusion of the Euro Area and U.K. real exchange rates")
(terpri)
(princ "F-Statistic = ")(princ FStat)(terpri)
(princ " P-Value = ")(princ PV)(terpri)
((terpri)

Notice that our code uses USSE and udf, the sum of squared errors and degrees
of freedom from the unrestricted regression. The results are as follows.

ORDINARY LEAST SQUARES REGRESSION

Dependent Variable: LREXCAUS  ---  Time Period 1999Q1 -- 2010Q4

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-stat</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.258</td>
<td>0.323</td>
<td>-0.799</td>
</tr>
<tr>
<td>LRPCOMXEN</td>
<td>0.604</td>
<td>0.071</td>
<td>8.503</td>
</tr>
<tr>
<td>LRPENERGY</td>
<td>0.243</td>
<td>0.019</td>
<td>12.833</td>
</tr>
<tr>
<td>DNCICAUS</td>
<td>0.021</td>
<td>0.002</td>
<td>12.153</td>
</tr>
</tbody>
</table>

Number of Observations: 48
Degrees of Freedom: 44
R-Squared: 0.9601108944743519
Adjusted R-Squared: 0.957391182733967
Sum of Squared Errors: 0.04959814935364273
F-Statistic: 353.0193587127761
   P-Value 0.0

43
F-Test for exclusion of the Euro Area, and U.K. real exchange rates
F-Statistic = 4.591693136577879
P-Value = 0.01572265112925597

BOOTSTRAPPED COEFFICIENTS

<table>
<thead>
<tr>
<th>Quantiles</th>
<th>Constant</th>
<th>LRPCOMXEN</th>
<th>LRPENERGY</th>
<th>DNCICAUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>.01</td>
<td>-1.0846</td>
<td>0.4656</td>
<td>0.1980</td>
<td>0.0185</td>
</tr>
<tr>
<td>.025</td>
<td>-0.9457</td>
<td>0.4951</td>
<td>0.2051</td>
<td>0.0188</td>
</tr>
<tr>
<td>.05</td>
<td>-0.8617</td>
<td>0.5175</td>
<td>0.2103</td>
<td>0.0192</td>
</tr>
<tr>
<td>.10</td>
<td>-0.7369</td>
<td>0.5406</td>
<td>0.2161</td>
<td>0.0196</td>
</tr>
<tr>
<td>.25</td>
<td>-0.5529</td>
<td>0.5826</td>
<td>0.2283</td>
<td>0.0204</td>
</tr>
<tr>
<td>.75</td>
<td>-0.1567</td>
<td>0.6720</td>
<td>0.2529</td>
<td>0.0226</td>
</tr>
<tr>
<td>.9</td>
<td>0.0423</td>
<td>0.7110</td>
<td>0.2627</td>
<td>0.0235</td>
</tr>
<tr>
<td>.95</td>
<td>0.1353</td>
<td>0.7379</td>
<td>0.2700</td>
<td>0.0242</td>
</tr>
<tr>
<td>.975</td>
<td>0.2448</td>
<td>0.7573</td>
<td>0.2771</td>
<td>0.0245</td>
</tr>
<tr>
<td>.99</td>
<td>0.2448</td>
<td>0.7927</td>
<td>0.2860</td>
<td>0.0252</td>
</tr>
<tr>
<td>median</td>
<td>-0.3543</td>
<td>0.6247</td>
<td>0.2403</td>
<td>0.0215</td>
</tr>
<tr>
<td>mean</td>
<td>-0.3507</td>
<td>0.6256</td>
<td>0.2403</td>
<td>0.0216</td>
</tr>
</tbody>
</table>

We can clearly reject the null-hypothesis that the Euro Area and U.K. real exchange rates with respect the the U.S. together contribute nothing to the explanation of the Canadian real exchange rate with respect to the United States. Again, the bootstrapped range of the real net capital inflow coefficient in this restricted regression does not include the coefficient value obtained in the regression for the whole period. An F-test of the restriction of excluding the Japanese and U.K. real exchange rates from that regression was unnecessary because the two variables were not statistically significant when included individually.

In general, we have to conclude that little could be added to these real exchange rate results and the analysis that follows by using an instrumental variables approach. That does not mean, of course, that problems of simultaneity can be ignored. We handle them here by developing a mathematical model and incorporating into it values of parameters estimated elsewhere.
Despite the statistical significance of the log real GDP variables and the unemployment rates in the long-period regression above, there is a remarkable similarity of the business cycles of not only Canada and the United States but other industrial countries as well. In addition, there is substantial evidence indicating very little difference between Canadian monetary policy and that in the United States.\textsuperscript{11} In fact, even the underlying core inflation rates in the two countries are pretty much the same. This policy makes sense for Canada because the two economies are closely interrelated and the Bank of Canada hires its economists from the same pool as does the U.S. Federal Reserve Bank. Also, the benefits of counter-cyclical policy actions in a big country like the United States are not entirely clear, raising even more serious doubts about the benefits of counter-cyclical deviations of Canadian from U.S. monetary policy. Should the economists at the Bank of Canada disagree with what the U.S. authorities are doing, the most important policy tool at their disposal is probably the telephone. There is no requirement, however, that the Canadian authorities should adopt the U.S. inflation target although it is nevertheless reasonable to expect that the inflation-control objectives of the two countries will be the same.

Suppose that, contrary to wisdom and the above evidence, the Bank of Canada decides to conduct a substantial expansionary monetary policy.\textsuperscript{12} Given the fact that Canadian interest rates are determined in the world market, the effect of such a policy will be to devalue the nominal exchange rate and, given the short-run rigidity of prices, the real exchange rate as well.

A problem here is the possibility of exchange rate overshooting if the Canadian authorities operate directly on the monetary aggregates. This can be seen from the fact that the demand for nominal money holdings can be expressed as

\[
M = P_N^\theta (\tilde{P}_T / \Pi)^{1-\theta} L(\tilde{r} + E_p, Y)
\]  

where \(M\) is the nominal money stock, \(P_N\) is the price of the non-traded component of domestic output, \(\Pi\) is the nominal exchange rate—that is, the

\textsuperscript{11}For a presentation of this evidence, see Chapter 11 of my book referred to previously.

\textsuperscript{12}A more realistic situation, in which it would want to do the opposite, would occur if for some reason the Canadian core inflation rate became much too high relative to the core inflation rates in the U.S. and abroad and above the rate desired by Canadian residents. A situation where the Canadian inflation rate is too low, requiring an expansionary monetary policy, is hard to imagine, although that situation has occurred in Japan.
foreign currency price of domestic currency—and $\tilde{P}_T$ is the foreign currency price of the traded component of domestic output. The domestic price of the traded component of domestic output is thus $\tilde{P}_T/\Pi$. The world and domestic real interest rate is denoted by $\tilde{r}$ and the expected rate of domestic inflation is $\tilde{E}_p$, making $\tilde{r} + \tilde{E}_p$ is the domestic nominal interest rate, and $Y$ is the level of domestic real output. The function $L(\tilde{r} + \tilde{E}_p, Y)$ is the demand function for domestic real money balances. An increase in $M$ on the left side of the above equation must be matched by and equal increase in the right side for money market equilibrium to be maintained. Since output prices cannot change in the short-run, the fall in $\Pi$ will reduce the real exchange rate and shift world demand onto domestic output, causing $Y$ to increase by an amount which, together with the exchange rate induced rise in the domestic price level, will increase right side of the equation to equal the increase in the left side. The problem is that it takes time for the balance of trade and the level of real income to respond to a fall in the real exchange rate. Until a response can take place the only equilibrating mechanism is a fall in the nominal exchange rate and corresponding rise in the domestic prices of the traded components of output. Taking the logarithm of equation (16) under these circumstances where prices, the real and nominal interest rates and real output are constant, we obtain

\[
\log(M) = \theta \log(P_N) + (1 - \theta)[\log(\tilde{P}_T) - \log(\Pi)] + \log(L(\tilde{r} + \tilde{E}_p, Y))
\]

\[
= -(1 - \theta) \log(\Pi)
\]

which can be rewritten as

\[
\log(\Pi) = -\frac{1}{1 - \theta} \log(M) .
\]

Taking the total differential of the above expression yields

\[
\frac{d\Pi}{\Pi} = -\frac{1}{1 - \theta} \frac{dM}{M} .
\]

Under the reasonable assumption that two-thirds of domestic output is non-traded, a one percent increase in the nominal money supply will cause the nominal exchange rate to fall by three percent. And this ignores the fact that it will take a period of time for the domestic currency prices of the traded components of output to adjust to the exchange rate change. In the days during which these prices remain unchanged the nominal exchange
rate may fall much further, with the only equilibrating mechanism being a speculative one arising from knowledge that there has to be a lower limit to the long-run equilibrium level of the exchange rate, speculative bubbles notwithstanding. The last thing a central bank wants to be accused of is having created disorderly markets.

Since there is no prospect of manipulating underlying domestic real interest rates, given a world-wide capital market, the only real control option available to the domestic authorities will be downward pressure on the nominal exchange rate. There are nevertheless two reasons why it makes good sense for the Bank of Canada to announce targets for, and exercise control over, the interest rate at which it will lend reserves to the domestic banking system. First, setting a target for the overnight lending rate helps establish public awareness of the Bank’s commitment to its inflation target. It is extremely important that the public have an appropriate expected inflation rate because by following an orderly markets monetary policy—that is, keeping the nominal exchange rate from jumping sharply outside normal trading ranges—the Bank of Canada can end up financing that expected inflation rate. Second, by controlling the rate at which it will lend to the banking system and at which the commercial banks will be able to borrow from each other, the Bank of Canada can affect the profitability to commercial banks of expanding there reserves and deposits and thereby exercise an element of direct influence on money supply growth followed by hopefully gradual real and nominal exchange rate changes.

In the long-run, of course, the expansion of output and employment resulting from the devaluation of the Canadian dollar and fall in the real exchange rate that will inevitably result from increased domestic monetary expansion will lead to upward pressure on and increases in the Canadian price level that will reduce domestic output and raise the real exchange rate back to their full-employment levels.

Our purpose now is to try to determine empirically the effect of a fall in the real exchange rate due to monetary policy on output and employment and the current account balance and net capital flow—that is, to empirically account for and measure the shifts and slope of the BT line along with shifts of the SI line resulting from monetary policy induced short-run real exchange rate changes. Since considerable simultaneity is involved, the best approach is to write down equations representing the determinants of BT and SI and then see what can be accomplished using the estimated coefficients in our
regression above, incorporating other available information as necessary. The equation of the BT curve can be written as

\[ q = \alpha + \beta \hat{B}_T + \gamma E \]  

(20)

where \( q \) is the logarithm of the real exchange rate, \( \beta < 0 \) is the slope of the BT curve, \( \hat{B}_T \) is the full-employment current account balance and associated full-employment net capital outflow as a percentage of the current full-employment level of domestic GDP under the assumption that the U.S. net capital inflow and full-employment GDP are unchanged and therefore incorporated in the constant term \( \alpha \). Finally, \( E \) is the Canadian employment rate (percentage of the labour force employed), with the U.S. unemployment rate being constant and also incorporated in \( \alpha \), and \( \gamma < 0 \) is the change in the log of the real exchange rate in response to a change in the percentage of the labour force employed. According to our regression result above, \( \beta = -0.032 \), and \( \gamma = -0.038 \). The magnitude of \( \alpha \) will not be of interest in the analysis that follows. The actual current account balance, which we call \( B_T \), is equal to the full-employment current account balance minus any increase in imports that results from a subsequent change in the employment rate.

\[ B_T = \hat{B}_T - mY = \hat{B}_T - m\delta E \]  

(21)

where \( m \) is the marginal propensity to import out of a change in current income \( Y \) and \( \delta \) is the increase in that income, as a percentage of its full employment level, produced by a one percentage point expansion of the level of employment. Finally, we must impose the fact that savings minus investment under less-than-full-employment conditions must equal the less-than-full-employment current account balance—that is,

\[ B_T = sY - I = s\delta E - I \]  

(22)

where \( s \) is the marginal propensity to save out of the change in current income and \( I \) is a constant equal to the constant underlying level of domestic investment where we keep in mind that, by construction, the real interest rate is unchanged and the level of employment and savings and investment do not change abroad. Equations (21) and (22) together yield the following expression for \( \hat{B}_T \),

\[ \hat{B}_T = (m + s)\delta E - I \]  

(23)
which upon substitution into (20) yields

\[ q = \alpha + \beta (m + s) \delta \log E - \beta I + \gamma E \]

\[ = \alpha - \beta I + [\beta (m + s) \delta + \gamma] E. \]  

(24)

Rearrangement of the above equation produces the response of the employment rate to a monetary policy induced change in \( q \).

\[ E = A + \frac{1}{\beta (m + s) \delta + \gamma} q \]  

(25)

where

\[ A = \frac{\alpha - \beta I}{\beta (m + s) \delta + \gamma} \]

is a constant term that incorporates exogenous shifts of the BT curve on account of forces beyond our present concern.

If the public fully understands what is happening and inter-temporally smooths consumption, the entire transitory shock to income will be saved, so \( s \) will equal unity and \( m \) will equal zero. We then need only to specify a value for \( \delta \), which represents the effect of a one percentage point increase in the percentage of the labour force employed on the logarithm of the level of output. If the aggregate production function is Cobb-Douglas, this will equal the share of labour in output. There is controversy over the exact magnitude of this share because of the complexities of trying to estimate it from available data. A rough guess would postulate a share of 0.7. However, in the present circumstances where there is a variation of the level of employment of a given stock of labour under given technological conditions the possibility arises that the utilization of capital could change in a different way than it would under full-employment conditions—in particular, the current level of the capital stock could become over- or under-employed in the same way as labour. This would suggest a higher value for \( \delta \). In addition, of course, the elasticity of substitution of labour for capital in the full-employment situation may be different from the Cobb-Douglas value of unity. Under these circumstances it might be useful to attempt to calculate \( \delta \) by regressing the level of output on the level of employment. First, we graph the time paths of the percentage deviation of Canadian real GDP from its trend level, along with the percentage of the labour force employed. This graph is presented on the next page.
The two series are obviously correlated and it is clear that the variation of real GDP around its trend is much greater than the variation of the percentage of the labour force employed. When we regress the logarithm of Canadian real GDP on the percentage of the labour force employed together with trend using our OLS function we obtain the following result.
ORDINARY LEAST SQUARES REGRESSION

Dependent Variable: LOG of RGDPCA  ---  Time Period 1976Q1 -- 2010Q4

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-stat</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>2.574</td>
<td>0.173</td>
<td>14.859</td>
</tr>
<tr>
<td>Trend</td>
<td>0.007</td>
<td>0.000</td>
<td>57.899</td>
</tr>
<tr>
<td>EMPRCA</td>
<td>0.014</td>
<td>0.002</td>
<td>7.277</td>
</tr>
</tbody>
</table>

Number of Observations: 140
Degrees of Freedom: 137
R-Squared: 0.9942048936116979
Adjusted R-Squared: 0.9941202935184381
Sum of Squared Errors: 0.06131114939203919
F-Statistic: 11751.817938989792
P-Value 0.0

Coefficient Standard Errors are Newey-West HAC Adjusted with Lag = 3

Breusch-Pagan ChiSquare Statistic: 1.902153702969665
P-Value 0.38632478496289024

LM-Test for Serial Correlation of Residuals:
Number of Lags = 3
Chisquare Statistic = 1358.4379478207427
P-Value = 0.0

Had we used the percentage deviation of real GDP from its trend and left out the trend variable, the result would be the same except for the magnitude of the constant term. The fit is remarkably good but there is clearly substantial serial correlation in the residuals. Indeed, the residuals, when plotted against time, are essentially equal to the unemployment rate over the percentage deviation of real GDP variable from its trend in the graph above. The presence of serial correlation in the residuals indicates that variables affecting real GDP are left out of the regression. Since we have no way of knowing what these variables are there is no way of using instrumental variables to correct the simultaneity bias problem. The coefficient of EMPRCA
implies that an increase in the level of employment of one percent will lead to an increase in real GDP of .14 percent. This makes no sense, so we have to rely on our knowledge of the share of labour, possibly allowing for some increase in the utilization of capital to establish a value for $\delta$. Allowing for the possibility that increased capital utilization could accompany short-run employment expansion, a value for $\delta$ of around 0.7 would seem reasonable and an interval of 0.6 through 0.75 would seem a reasonable range to allow for error.\footnote{I would like to thank my colleague Margarida Duarte for a helpful discussion of the range of possible values.}

Assuming that a one-percent increase in the level of employment increases output by .7 percent of its full-employment level so that $\delta = .7$, and using the values of $\beta$ and $\gamma$ implied by our regression result above, the total differential of (25) implies that

$$dE = -\frac{1}{(.032)(.7) + .038} dq = \frac{1}{.0604} dq = -16.33 dq$$

which implies that a one percent or .01 fall in the logarithm of the real exchange rate will lead to a percentage increase in the level of employment of a bit more than 0.16 percent. And to increase the level of employment by 1 percentage point, the Bank of Canada would have to expand the money supply sufficiently to reduce the nominal and real exchange rate by about 6 percent. In the case where $\delta = .75$, the above magnitudes change to .062 and 16.13 and when we let $\delta = .6$, they become .057 and 17.48. So the range of real exchange rate devaluation required to increase the level of employment by 1 percentage point would be from 5.7 to 6.2 percent.

Our assumption that the public regards the increase in current income as entirely transitory is probably unrealistic, given the lack of current information about the cause of the observed increase in its income. Suppose, to take the most extreme case, that the public incorrectly regards the observed increase in its income as permanent. The data analyzed in the Excel spreadsheet file conincca.xls suggest that an appropriate value for the fraction of permanent income saved is .213 and for the marginal propensity to import .259. The the above expression then yields

$$dE = -\frac{1}{(.032)(.213 + .259)(.7) + .038} dq = \frac{1}{.0486} dq = -20.58 dq$$
and a monetary expansion induced fall in the exchange rate of slightly under 5 percent would be required to increase the level of employment by 1 percent of the labour force. A smaller fall in the exchange rate and less monetary expansion is required because of the multiplier effect of expansion of consumption expenditure resulting from the increase in employment and income. When $\delta$ ranges between .6 and .75 the required fall in the exchange rate ranges between 4.71 and 4.94 percent.

Overall, the truth probably lies somewhere within the range of 4.7 and 6.2 percent. As a rough guess we might conclude that to get a one percent increase in the fraction of the labour force employed, a monetary expansion sufficient to reduce the nominal (and real) exchange rate by between 5 and 6 percent will probably be required.

**Figure 11: Real Exchange Rate Determination**

This result can be seen graphically in Figure 11 above. A fall in the real exchange rate from a to b will, because of the simultaneity issues involved, increase the current account balance by less than c d for one definite reason plus possibly another plausible reason. First, the associated short-run increase in employment will increase domestic relative to foreign output, reducing its value in world markets and shifting the BT curve downward to the left. In addition, it is quite likely that domestic residents will not realize...
that the observed increase in output and income is entirely transitory, with the result that consumption and imports will increase, shifting the BT curve downward to the left by an additional amount. The increase in the current account balance, and in savings relative to investment will thus be of a magnitude like $c_e$. And in this event, although the current account balance will increase by less, the level of employment and income will increase by more as a consequence of the multiplier effect of the increase in consumption.

Finally, the fact that a monetary expansion induced fall in the real exchange rate of between five and six percent will increase domestic employment by one percentage point does not mean that the Bank of Canada should try to use nominal exchange rate manipulation to achieve less variability of the Canadian unemployment rate. First of all, as the Bank of Canada begins moving the nominal and real exchange rates it loses sight of their equilibrium level and, hence, the magnitude of the effect of its policy. Second, the Bank of Canada observes the unemployment rate and current output with a substantial lag, and the effects of its change in the real exchange rate on output and employment will only take place after a further lag. As a result, the expansion of employment induced by Bank policy may well begin to occur just as the economy is recovering from the recession and thereby accentuate subsequent inflationary pressure. Monetary manipulation of the nominal and real exchange rates is a useful policy only under the circumstances in which the Bank of Canada needs to bring about a significant change in the underlying domestic core inflation rate. Given its conduct of monetary policy in the past, a situation of requiring this kind of action is very unlikely to occur.
Figure 11: The net capital inflows of Canada and the United States as percentages of their GDPs and the excess of Canadian over United States percentages.

We have been analyzing the excess of the net capital inflow into Canada as a percentage of that country’s GDP over the net capital inflow into the United States as a percentage of U.S. GDP. That focus was based on ease of exposition. As can be seen from the figure above, the United States experienced a net capital inflow for all but the early years of the period examined and Canada experienced net capital outflows (negative net capital inflows) for all but three small intervals over the period.