IS CANADA'S AGGREGATE PRODUCTION FUNCTION COBB-DOUGLAS? ESTIMATION OF THE ELASTICITY OF SUBSTITUTION BETWEEN CAPITAL AND LABOR

by

Haitao Xiang

Bachelor of Economics & Bachelor of Law
Huazhong University of Science and Technology, 1996
Master of Economics
Zhongnan University of Economics and Law, 2002

PROJECT
SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

In the Department
of
Economics

SIMON FRASER UNIVERSITY

July 2004

APPROVAL

Name: **Haitao Xiang**

Degree: **M. A. (Economics)**

Title of Project: **Is Canada's Aggregate Production Function Cobb- Douglas? Estimation Of The Elasticity Of Substitution Between Capital And Labor**

Examining Committee:

Chair: Lawrence Boland

Ken Kasa
Senior Supervisor

Peter Kennedy
Supervisor

Brian Krauth
Internal Examiner

Date Approved: Friday, July 30, 2004

Abstract

Based on U.S. data, Antràs (2004) illustrates that the assumption of Hicks-neutral technological change necessarily biases estimates of the elasticity of substitution between capital and labor towards one, given Berndt's (1976) specification. Modifying the specification to allow for biased technological change, he obtains significantly lower estimates of the elasticity. Using Canadian data, I obtain similar results. This suggests that a Cobb-Douglas specification of Canada's aggregate production function may be misleading.

Acknowledgements

I would like to express my sincere thanks to Professor Kenneth Kasa, under whose supervision I could complete this project. I am also very thankful to Professor Peter Kennedy and Professor Brian Krauth for their valuable suggestions. I should not forget to thank Professor Richard Harris, who offered me the opportunity to write on this topic. My appreciation is also extended to Professor Pol Antràs of Harvard University, for his great help with my replication of his inspiring paper.

I would like to thank my wife, Qiong, and my families, for their support and encouragement during the past years. I would also like to thank all the faculty and staff in the Department of Economics, and all the friends I have met here at Simon Fraser University.

Table of Contents

List of Tables and Figures

1 Introduction

The elasticity of substitution between capital and labor plays an important role in economic theory. In particular, different values of this elasticity result in different implications in models of economic growth or income distribution. For instance, the sustainability of long-run growth in the absence of technological change depends crucially on whether the elasticity of substitution is greater than or smaller than one. In addition, the elasticity of substitution may affect tax incidence between capital and labor. A lower elasticity is also associated with a lower response of investment to tax benefits. Therefore, the elasticity of substitution between capital and labor is of significant policy interest[1].

Empirically, the value of the elasticity of substitution for the U.S. manufacturing sector has been widely estimated since the derivation of the Constant Elasticity of Substitution (CES) production function by Arrow et al. (1961). In particular, Berndt (1976) reconciled the difference between time-series and cross-sectional studies, by use of high-quality data, and found the estimates to be insignificantly different from one. As is well known, a CES production function with unity elasticity of substitution is Cobb-Douglas.

Antràs (2004) demonstrates that the assumption of Hicks-neutral technological change necessarily biases estimates of the elasticity of substitution between capital and labor towards one, given Berndt's (1976) specification. The source of the bias can be understood as follows.

[1] See Antràs (2004) for the literature sources of more examples concerning implications of the elasticity of substitution.

Suppose aggregate output Y_t can be represented by a production function of the form:

$Y_t = A_t F(K_t, L_t)$,

which is characterized by constant returns to scale in the two inputs, capital K_t and labor L_t. The parameter A_t is an index of technological efficiency, in the sense that it has no effect on the ratio of marginal products for a given capital-labor ratio. Profit maximization by firms in a competitive framework delivers two optimality marginal product conditions: $r_t = f'(k_t)$ and $w_t = f(k_t) - f'(k_t)k_t$, where k is the capital-labor ratio, $f(k)$ is output per unit of labor, and r and w are the real rental prices of capital and labor, respectively. Combining these two conditions, we have

$$\frac{r_t}{w_t} \cdot k_t = \frac{f'(k_t)k_t}{f(k_t) - f'(k_t)k_t}, \text{ or}$$

$$\frac{r_t K_t}{w_t L_t} = \frac{f'(k_t)k_t / f(k_t)}{1 - f'(k_t)k_t / f(k_t)}.$$

The left-hand side of this equation is the ratio of capital and labor income, which is known to be stable in the United States, while the capital-labor ratio has steadily increased for the post-war years.[2] Then this equation can be consistent with the U.S. data only if $f'(k_t)k_t / f(k_t)$ is not a function of k_t, i.e., $f'(k_t)k_t / f(k_t) = \alpha$ where α is a constant. Solving this differential equation yields $f(k_t) = Ck_t^{\alpha}$ (C is a constant of integration), which is a Cobb-Douglas production function, characterized by a unit elasticity of substitution. Since the approach of Berndt (1976) consists of running log-linear specifications closely

[2] This is also true for the Canadian economy.

related to the expression above, his finding of a unit elasticity of substitution should not be surprising.

As noted in Antràs (2004), the Cobb-Douglas production function is no longer the only one consistent with stable factor shares, when technological change is allowed to affect the ratio of marginal products, i.e., the technological change is not Hicks-neutral. Modifying the specification to allow for biased technological change, he obtained significantly lower estimates of the elasticity.

The estimation of Antràs (2004) is based on time-series data from the private sector of the U.S. economy for the period 1948-1998. To verify his arguments, I make a replication of that paper with Canadian data. And the results of my estimation turn out to be consistent with those obtained by Antràs: With Hicks-neutral technological change, more than half of the estimates are insignificantly different from one, while most of them are significantly lower than one under biased technological change.

2 Model Specification

Following Antràs (2004), I assume that aggregate production in the Canadian business sector can be represented by a constant returns to scale production function, which is characterized by a constant elasticity of substitution between the two factors, capital and labor. Arrow et al. (1961) derived a CES functional form for the production function as below:

$$Y_t = A_t\left[\delta K_t^{\frac{\sigma-1}{\sigma}} + (1-\delta)L_t^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}},$$

where Y_t is real output, K_t is the flow of capital services, L_t is the flow of labor services, A_t is a Hicks-neutral technological shifter, δ is a distribution parameter, and the constant σ is the elasticity of substitution between capital and labor. σ is defined as $\sigma = d\log(K/L)/d\log(F_L/F_K)$, where F_L and F_K are the marginal products of capital and labor, respectively. Berndt (1976) defines the aggregate input function $F_t \equiv Y_t / A_t$, which is independent of A_t given the assumption of Hicks-neutral technological change.

Assuming competitive markets, profit maximization by firms gives two first-order marginal productivity relations: real factor prices equal real value of their marginal products. Taking logarithms and adding an error term, we obtain

$$\log(F_t / K_t) = \alpha_1 + \sigma\log(R_t / P_t) + \varepsilon_{1,t} \quad (1)$$

$$\log(F_t / L_t) = \alpha_2 + \sigma\log(W_t / P_t) + \varepsilon_{2,t} \quad (2)$$

where R_t, W_t, and P_t are the prices of capital services, labor services, and aggregate input F_t, respectively, and α_1 and α_2 are constants that depend on σ and δ. As pointed out by Antràs, the disturbance terms can be explained as optimization errors. If we subtract (1) from (2), we obtain a third equation

$$\log(K_t / L_t) = \alpha_3 + \sigma \log(W_t / R_t) + \varepsilon_{3,t} \tag{3}$$

In the above equations prices are assumed to be exogenous. However, if actually they are endogenous, OLS estimates will be inconsistent with the direction of the bias unknown. As noted by Berndt (1976), even if both prices and quantities are endogenous, in our small sample case it may be desirable to ignore the simultaneous equations bias and estimate the reciprocal regressions from equations (1) through (3) by OLS:

$$\log(R_t / P_t) = \alpha_4 + (1/\sigma) \log(F_t / K_t) + \varepsilon_{4,t} \tag{4}$$

$$\log(W_t / P_t) = \alpha_5 + (1/\sigma) \log(F_t / L_t) + \varepsilon_{5,t} \tag{5}$$

$$\log(W_t / R_t) = \alpha_6 + (1/\sigma) \log(K_t / L_t) + \varepsilon_{6,t} \tag{6}$$

To be convenient, denote the estimates of σ based on equations (1) through (6) by σ_i, i = 1,…, 6. As shown in Berndt (1976), in this bivariate regression setting, we have following relationships for the OLS estimates:

$$\frac{\sigma_1}{\sigma_4} = R_1^2 = R_4^2; \; \frac{\sigma_2}{\sigma_5} = R_2^2 = R_5^2; \; \frac{\sigma_3}{\sigma_6} = R_3^2 = R_6^2,$$

where R_i^2 refers to the R-square in equation i, for R^2 here is just the squared sample correlation coefficient between the two variables. Since $R^2 \leq 1$, we always have inequalities $\sigma_1 \le \sigma_4$, $\sigma_2 \le \sigma_5$, $\sigma_3 \le \sigma_6$. And it is clear that the standard and reverse estimates are closer if the R-square in the OLS regression is larger.

Berndt (1976) pointed out that estimates of the elasticity of substitution seem to vary systematically with the choice of functional form: regressions based on the marginal product of capital equation (1) generally produce lower estimates of σ than regressions based on the marginal product of labor equation (2). And it is confirmed again by Antràs (2004) – his estimates are consistent with this empirical regularity: $\sigma_1 < \sigma_2$.

3 Data Construction and Sources

As can be seen from equation (1) through (6), the data required for the estimation are: the flow of labor services L_t, the nominal price of labor services W_t, the flow of capital services K_t, the rental price of capital R_t, the aggregate input index F_t, and the price of aggregate input P_t.

Antràs (2004) used the U.S. private sector data. However, there is no such a counterpart in Canadian data. Statistics Canada segments the economy into business sector and non-business sector, instead of private sector and public sector. Since the differences are not expected to be significant (see Harchaoui et al., 2001, pp. 168-169), I use Canadian business sector data for estimation.

For my project, all the data are retrieved from CANSIM database of Statistics Canada. Thanks to the Statistics Canada's Multifactor Productivity Program, I could directly obtain annual Fisher chained index series for all the six variables for the period 1961-1997.

L_t, the flow of labor services, or labor input, is measured as a weighted sum of hours worked by industry where the weighs are defined as the industry's share in the total labor compensation. The weights are assumed to reflect the differences in the composition of the labor force, or labor quality, by industry.

W_t, the nominal price of labor services, or price of labor input, is taken to be equal the total compensation of all jobs divided by L_t. Compensation of all jobs includes wages, salaries and supplementary labor income accrued to employees and self-employed.

K_t, the flow of capital services, or capital input, is assumed to be proportional to real capital stock, which is weighted by asset prices.

Total capital income is a residual as total income minus all other input costs. And the rental price of capital, R_t, is computed as the ratio of total capital income to the real capital stock K_t.

The aggregate input index F_t is calculated as follows:

$$\frac{F_t}{F_{t-1}} = \frac{1}{2}\left(s_t^L + s_{t-1}^L\right)\frac{L_t}{L_{t-1}} + \frac{1}{2}\left(s_t^K + s_{t-1}^K\right)\frac{K_t}{K_{t-1}},$$

where s_t^L and s_t^K represent the input shares in terms of compensation (hence $s_t^L + s_t^K = 1$).

The price of combined input index P_t is constructed implicitly as $(R_t K_t + W_t L_t)/F_t$, i.e., the sum of labor compensation and capital compensation divided by F_t.

As shown by Berndt (1976), the quality of data is essential to the estimation. Guided in this way, Antràs (2004) illustrated the effect of data quality on the estimates of the elasticity, by experimenting with different methods in the construction of these variables. Antràs was able to benefit from recent literature on improvement of measurement of the U.S. capital and labor input. Statistics Canada is also proceeding with revision of estimates of capital and labor input, to reflect the heterogeneity and changing compositions of both types of input. The methodology used for this revision can be found in Gu et al. (2003) and Harchaoui et al. (2003). Unfortunately, for the time being the available quality-adjusted data only covers the period 1981-2000. If we try to do estimation based on these refined data, we have to face serious small-sample bias. However, I still present the results from the estimation based on this sample of only 20 observations. For this data set, I use labor input and capital input volume index, which are available as CANSIM series, as L_t and K_t. And I construct their associated price index as the ratio of factor compensation and input index. Following Berndt (1976) and Antràs (2004), I construct the price of aggregate input P_t as a Tornqvist aggregate index of W_t and R_t. And the aggregate input index F_t is computed implicitly as $(R_t K_t + W_t L_t) / P_t$.

Hereafter I denote the quality-adjusted data as data A and the unadjusted data as data B.

4 Estimates under Hicks-Neutral Technological Change

As same as Antràs (2004), I present estimates of the elasticity of substitution between capital and labor based on different estimation methods. First I report simple ordinary least squares estimates of equations (1) through (6) for both data configurations. Then I try to solve the problems caused by serial correlation in disturbances, endogeneity of the regressors, and nonstationarity of the series.

4.1 Least Squares Estimation

Table 1 presents OLS estimates of equations (1) through (6) for both data sets. Most notable is the huge difference between the estimates from different data configurations. For data configuration B, all the estimates of the elasticity are very close to unity. But data set A produces estimates quite lower than one. When we compare the standard errors and R-squares, we can find that quality-adjusted data make a poorer estimation with higher standard errors and lower R-squares, which are opposite to the finding of Antràs (2004). As high R-square makes standard and reciprocal estimates from the same first-order condition approach to each other, the estimates in column II cluster much closer, ranging from 0.997 to 1.156. In contrast, the estimates of column I have a wider range from 0.359 to 0.742. The null hypothesis of a unit elasticity of substitution cannot be rejected at the 5% significance level for equations (1) through (4) for data set B, and equations (4) through (6) for data set A. Another point to be noted is that the low Durbin-Watson statistics for all the specifications suggest existence of serial correlation in the residuals.

It seems that the inherent bias toward unity with assumption of Hicks-neutral technological change exposited by Antràs (2004) is also found with Canadian data. However, the effect of data quality on estimation is found to be totally different. The undesirable estimates out of quality-adjusted data are very likely to be caused by the small-sample bias. Therefore I will proceed with data configuration B only.

4.2 Feasible Generalized Least Squares Estimation

The low Durbin-Watson statistics in OLS estimation suggest that there exist serial correlation in the residuals. A Ljung-Box test for autocorrelation was performed for all six specifications, and the null hypothesis of no autocorrelation up to order k was rejected in all six regressions for all $k \leq 16$. Following Antràs (2004), I also assume a standard AR(1) process for the autocorrelation structure, i.e., $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$ where u_t is white noise.

To be consistent with Antràs (2004), I also applied the two-step Prais-Winsten procedure to data configuration B. Column III of Table 1 reports the FGLS estimates of the elasticity of substitution. All the FGLS estimates are lower than the OLS ones. The estimates from equations (1) through (3) are lower than 0.80, far away from one. But the estimates from equations (4) through (6) become closer to one, ranging from 0.974 to 1.006. Similar to the results of Antràs (2004), the FGLS standard errors are higher than the OLS ones but the difference is not so big. Now the estimates σ_4, σ_5, and σ_6 are insignificantly different from one, while the null hypothesis of unity elasticity is rejected for the first three estimates.

Table 1 Estimates with Hicks-Neutral Technological Change

		OLS		FGLS	GIV
		A	B	B	B
		I	II	III	IV
1	σ_1	**0.376**	**0.997**	**0.760**	**0.672**
	S.E.	0.087	0.056	0.066	0.078
	R^2	0.511	0.902	0.792	0.687
	D-W	0.201	0.450	1.030	0.594
2	σ_2	**0.359**	**1.039**	**0.770**	**0.704**
	S.E.	0.087	0.059	0.069	0.086
	R^2	0.484	0.899	0.780	0.662
	D-W	0.205	0.447	0.983	0.641
3	σ_3	**0.370**	**1.013**	**0.763**	**0.680**
	S.E.	0.087	0.057	0.067	0.081
	R^2	0.501	0.901	0.788	0.677
	D-W	0.203	0.445	1.002	0.607
4	σ_4	**0.736**	**1.106**	**0.974**	**0.921**
	S.E.	0.170	0.062	0.081	0.112
	R^2	0.511	0.902	0.805	0.666
	D-W	0.491	0.525	1.253	1.238
5	σ_5	**0.742**	**1.156**	**1.006**	**0.979**
	S.E.	0.180	0.066	0.086	0.128
	R^2	0.484	0.899	0.795	0.632
	D-W	0.504	0.521	1.262	1.239
6	σ_6	**0.738**	**1.124**	**0.984**	**0.939**
	S.E.	0.174	0.063	0.083	0.117
	R^2	0.501	0.901	0.802	0.653
	D-W	0.496	0.519	1.243	1.231
No. of Obs.		20	37	37	36

4.3 Generalized Instrumental Variable Estimation

As explained in Antràs (2004), estimating equations (1) through (3) and their reverse specifications exposes the existence of an endogeneity problem. Antràs (2004) adopted three instruments as supply shifters to solve the problem of simultaneous equation bias: (1) U.S. population, (2) wages in the government sector, and (3) real capital stock owned by the government. To be consistent with his paper, I use similar instruments for my estimation: (1) Canadian population[3], (2) wages in the non-business sector[4], and (3) real capital stock owned by the government at the end of previous year[5]. The justification of the use of these three instruments and their construction methods are similar to those explained in Antràs's paper.

Instead of standard 2SLS procedure, a generalized instrumental variable (GIV) procedure developed by Fair (1970) is implemented in Antràs (2004), in order to solve the autocorrelation problem at the same time. Using the same estimation technique summarized in the Appendix A of Antràs (2004), I obtained GIV estimates of the elasticity of substitution, which are presented in Table 1 as column IV. For all the equations the GIV estimates are lower than the FGLS ones, while the GIV estimates σ_4, σ_5, and σ_6 are still close to one, ranging from 0.921 to 0.979. In Antràs (2004), the GIV estimation collapses all the estimates into the smallest interval around one (0.989, 1.017), among the three estimation procedures. However, it is not the same for my case. As same as the FGLS

[3] Only quarterly data are available. I use the first quarter data here.

[4] Wages in the non-business sector are computed as total compensation per job in the non-business sector deflated by aggregate input price index P_t.

[5] It is constructed as the nominal value of fix capital owned by the government divided by the implicit price index of government gross fixed capital formation.

estimates, the null hypothesis of a unit elasticity of substitution cannot be rejected only for the last three estimates, σ_4, σ_5 and σ_6.

Here another point should be noted. If we compare the estimates σ_1 and σ_2 (or σ_4 and σ_5) in columns II, III and IV, the latter is always larger. This is consistent with the empirical regularity found in the U.S. data, which is mentioned in Berndt (1976) and Antràs (2004): Regressions based on the marginal product of capital equation (1) generally produce lower estimates of σ than regressions based on the marginal product of labor equation (2).

4.4 Time Series Estimation

As a major difference from Berndt (1976), Antràs (2004) uses modern time series analysis to treat the nonstationarity nature of the data series.

Figure 1 Nostationarity of the Series

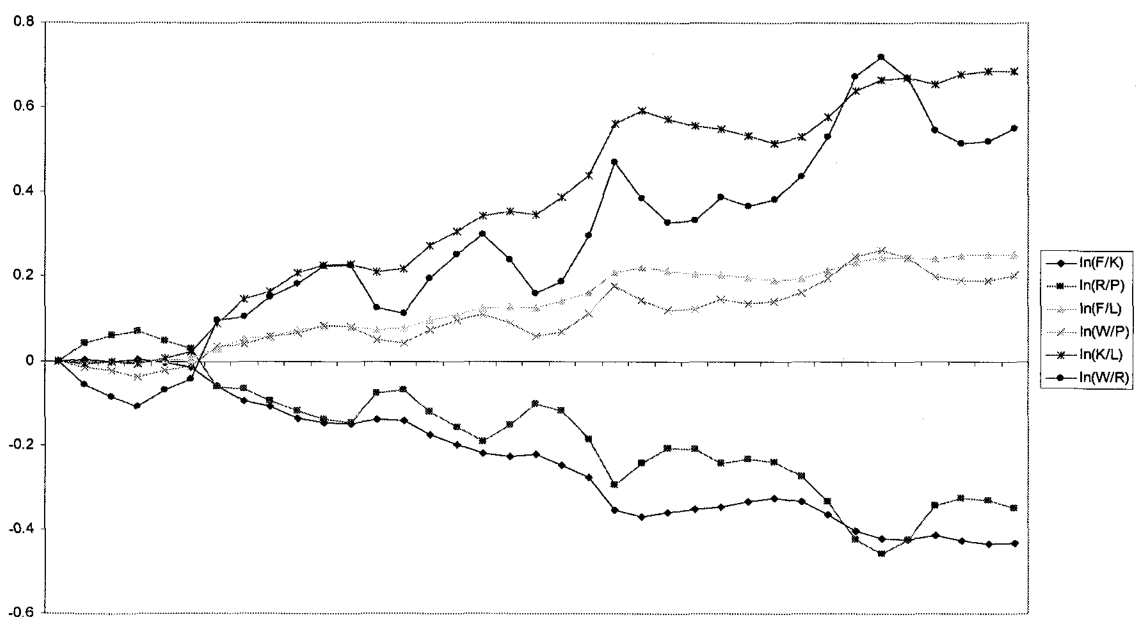


Figure 1 graphs the six series involved in our estimation, with the logarithms of the variables normalized to zero in 1961. As is similar to the U.S. version in Antràs (2004), two facts can be noticed from the figure. First, potential nonstationarity exists in each one of the series: the logarithms of W_t / P_t, F_t / L_t, W_t / R_t and K_t / L_t all trend upwards, whereas R_t / P_t and F_t / K_t show a downward trend. Second, the two variables in each of the specifications (1) through (6) follow similar trends. Therefore, we have a similar potential problem of spurious regression as noted in Antràs (2004). Although the high R-square and low Durbin-Watson statistics in our OLS estimation already imply this possibility, we still need formal test for the nonstationarity.

Table 2 reports the results of the unit root tests I performed on each of the series. The first row is the results of a Dickey-Fuller test of a unit root in the series. And the next two rows are results of Augmented Dickey-Fuller tests that include higher-order autoregressive terms to allow for serial correlation. The null hypothesis of a unit root cannot be rejected for each series whether we add in one or two lags. A Phillips-Perron test at truncation lags 2, 3 and 4 also gives the same result. The ADF and PP test are also performed on first difference of each of the six series. And the results show a rejection of the null hypothesis of a unit root in first differences, which means that all the series are not integrated of order two. Therefore, a conclusion can be drawn that all the series are integrated of order one. This implies that the OLS estimates are potentially subject to a spurious regression bias. OLS estimates are not consistent unless the two variables in each regression are cointegrated.

Table 2 Unit Root Tests

	$\log\left(\frac{F}{K}\right)$	$\log\left(\frac{R}{P}\right)$	$\log\left(\frac{F}{L}\right)$	$\log\left(\frac{W}{P}\right)$	$\log\left(\frac{K}{L}\right)$	$\log\left(\frac{W}{R}\right)$	5% Critical Value
ADF 0	-0.834	-0.922	-0.752	-0.962	-0.803	-0.934	-2.945
ADF 1	-1.256	-1.606	-1.215	-1.522	-1.242	-1.578	-2.947
ADF 2	-1.150	-1.268	-1.109	-1.197	-1.134	-1.239	-2.950
PP 2	-0.840	-1.043	-0.779	-1.075	-0.816	-1.053	-2.945
PP 3	-0.835	-0.936	-0.771	-0.957	-0.811	-0.942	-2.945
PP 4	-0.832	-0.834	-0.760	-0.858	-0.804	-0.841	-2.945
	$\Delta\log\left(\frac{F}{K}\right)$	$\Delta\log\left(\frac{R}{P}\right)$	$\Delta\log\left(\frac{F}{L}\right)$	$\Delta\log\left(\frac{W}{P}\right)$	$\Delta\log\left(\frac{K}{L}\right)$	$\Delta\log\left(\frac{W}{R}\right)$	5% Critical Value
ADF 1	-3.917	-4.951	-4.006	-5.048	-3.955	-4.995	-2.950
PP 3	-3.507	-4.078	-3.546	-3.967	-3.515	-4.016	-2.947

Table 3 presents the summary of two cointegration tests. Part A reports the results of Engle and Granger's (1987) residual-based Augmented Dickey-Fuller test. It is merely a ADF test on the residuals from the OLS regressions (1) through (6). However, the critical values of standard unit root tests are not appropriate. Mackinnon (1991) gives a set of parameters and a convenient formula to calculate critical values. I use the parameters for N=1 and "no trend" specification to calculate the critical values of 5% significant level, which are reported in the last column of part A. It is clear that the results are inconclusive. When the estimation includes one lagged first difference of the residuals, the null hypothesis of nonstationarity is rejected for all the six specifications, which means the two variables in the regression are cointegrated. However, the null hypothesis cannot be rejected for the same regressions if we include two lagged first differences, implying that the two variables in the regression are not integrated. This finding is the same as that reported in Antràs (2004).

Part B of Table 3 shows the results from the cointegration tests of Johansen and Juselius (1990), which have two versions. The max-lambda version tests the null hypothesis of the existence of r integrations against the alternative of $r+1$ cointegrations with $r = 0, 1, \ldots,$ k-1(k is the number of variables tested), while the trace version tests the null hypothesis of the existence of r integrations against the alternative of k cointegrations. To do this test, I chose the model with a constant and no trend, the same as that in Antràs (2004), for the cointegration equation. It can be seen that the null hypothesis of no integration cannot be rejected for all the six specifications, no matter one or two lags are included in the test. It seems to be safe to conclude that there is no integration for all the six specifications. However, this cointegration test requires a large sample size (about 300) to be reliable, according to Kennedy (2003, p 435). Considering our small sample size (only 37), it may be too risky to draw such a conclusion of no integration. And the inconclusive results presented by the residual-based ADF test can be regarded as some evidence of existence of possible cointegrations.

Table 3 Cointegration Tests

A. Residual-Based Augmented Dickey-Fuller Tests

	Residuals of eq(1)	Residuals of eq(2)	Residuals of eq(3)	Residuals of eq(4)	Residuals of eq(5)	Residuals of eq(6)	5% Critical Value
ADF 0	-2.277	-2.164	-2.221	-2.358	-2.287	-2.318	-2.945
ADF 1	-3.063	-3.093	-3.088	-3.407	-3.404	-3.421	-2.947
ADF 2	-2.295	-2.314	-2.296	-2.571	-2.557	-2.558	-2.950

B. Johansen-Juselius Cointegration Tests

	Max-Lambda				Trace			
Test	r =0 vs r =1		r ≤ 1 vs r =2		r =0 vs r =2		r ≤ 1 vs r =2	
No. of Lags	1	2	1	2	1	2	1	2
log(F/K) & log(R/P)	10.98	7.55	5.66	4.96	16.64	12.48	5.66	4.94
log(F/L) & log(W/P)	11.05	7.46	5.67	5.41	16.72	12.87	5.67	5.41
log(K/L) & log(W/R)	11.12	7.57	5.61	5.07	16.73	12.63	5.61	5.07
5% Critical Value	15.67		9.24		19.96		9.24	

Therefore, we would say that the cointegration tests are inconclusive, just like what Antràs (2004) did. Given this, our OLS estimates should be interpreted with caution because of a potential spurious regression bias. Antràs (2004) did not difference the data to solve the problem of spurious regression, since by doing so important long-run information would be lost and the interpretation of coefficients would become difficult. He also pointed out that the FGLS and GIV estimates are asymptotically equivalent to the estimates that would be obtained with the differenced data, given a unit root in the residuals of OLS regressions (Hamilton, 1994, p562). This implies that our estimates in columns III and IV of Table 1 are still consistent, at the cost of important long-run information in the data being neglected.

5 Estimates under Biased Technological Change

The results in the previous section provide some evidence in favour of a Cobb-Douglas specification of Canada's aggregate production function. All the six OLS estimates in column II in Table 1 are close to one, and the null hypothesis of a unit elasticity of substitution cannot be rejected for 10 out of the 18 estimates in columns II, III an IV. Therefore, our estimation based on Canadian data seems to be consistent with one of Antràs' (2004) conclusions: The assumption of Hicks-neutral technological change necessarily biases the estimates of elasticity of substitution towards one, which means an aggregate production function of Cobb-Douglas.

On the other hand, Antràs (2004) showed that in the presence of biased technological change, Berndt (1976) estimation equations are misspecified in a critical way, biasing the estimates towards finding the results that support a Cobb-Douglas production function. And he also offered some solutions to this misspecification problem. Modifying the econometric specification to allow for biased technical change, he obtained significantly lower estimates of the elasticity of substitution. In this section I will verify this point with Canadian data, using the same estimation procedure employed by Antràs (2004).

5.1 The Sources of the Bias

In Antràs (2004), the source of the bias is explained in the way as follows. Consider the Arrow et al. (1961) CES production function expanded to allow for non-neutral technological change

$$Y_t = \left[\delta \left(A_t^K K_t \right)^{\frac{\sigma-1}{\sigma}} + (1-\delta)\left(A_t^L L_t \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{7}$$

where A_t^K is an index of capital-augmenting efficiency and A_t^L is an index of labor-augmenting efficiency. In the presence of biased technological change, it is impossible to construct an index of aggregate input F_t that is independent of the efficiency indices A_t^K and A_t^L. In this case, the estimation equations (1), (2), (4) and (5), which include F_t and its associated price, are all misspecified in the sense that they suffer from an omitted-variable bias. If we take the first order conditions with respect to capital and labor, we can obtain the following expression, from which equations (3) and (6) are derived:

$$\log\left(\frac{W_t}{R_t}\right) = \log\left(\frac{1-\delta}{\delta}\right) + \left(\frac{1}{\sigma}\right)\log\left(\frac{K_t}{L_t}\right) + \left(\frac{1-\sigma}{\sigma}\right)\log\left(\frac{A_t^K}{A_t^L}\right) \tag{8}$$

Equation (8) shows that as long as $A_t^K \neq A_t^L$ (i.e., as long as technological change is non-neutral) equations (3) and (6) also suffer from an omitted-variable bias. Furthermore, we can subtract $\log(K_t / L_t)$ from equation (8) to obtain

$$\log\left(\frac{W_t L_t}{R_t K_t}\right) = \log\left(\frac{1-\delta}{\delta}\right) + \left(\frac{1-\sigma}{\sigma}\right)\log\left(\frac{A_t^K K_t}{A_t^L L_t}\right) \tag{9}$$

The left-hand side of equation (9) is the logarithm of the ratio of labor income and capital income, which is stable for the sample period. As noted in the introduction, the capital-labor ratio K_t / L_t on the right hand side has steadily increased during the same period. Therefore, if the bias in technological change is ignored, i.e., if the ratio A_t^K / A_t^L is not included in the regression, the estimate of $(1-\sigma)/\sigma$ will necessarily be close to zero, implying that the estimate of σ will necessarily be close to one. On the other hand, if K_t / L_t and A_t^L / A_t^K grow at the same rate, then steady factor shares can be consistent with any well-behaved production function, hence the elasticity of substitution may or may not be one.

5.2 Model Specification and Additional Data

As is clear now, when biased technological change is taken into account, it is impossible to construct an index of aggregate input F_t that is independent of the efficiency indices A_t^K and A_t^L. Therefore, it is necessary to control for these two efficiency indices in our estimation of elasticity of substitution. Following Antràs (2004), we assume that A_t^K and A_t^L grow at constant rates λ_K and λ_L. Then the production function becomes

$$Y_t = \left[\delta \left(A_0^K e^{\lambda_K \cdot t} K_t \right)^{\frac{\sigma-1}{\sigma}} + (1-\delta) \left(A_0^L e^{\lambda_L \cdot t} L_t \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} .$$

Now the first-order conditions for profit maximization become the following six specifications, analogous to equations (1) through (6) in section 2:

$$\log(Y_t / K_t) = \alpha_1' + \sigma \log(R_t / P_t^Y) + (1-\sigma)\lambda_K \cdot t + \varepsilon_{1,t} \tag{1'}$$

$$\log(Y_t / L_t) = \alpha_2' + \sigma \log(W_t / P_t^Y) + (1-\sigma)\lambda_L \cdot t + \varepsilon_{2,t} \tag{2'}$$

$$\log(K_t / L_t) = \alpha_3' + \sigma \log(W_t / R_t) + (1-\sigma)(\lambda_L - \lambda_K) \cdot t + \varepsilon_{3,t} \tag{3'}$$

$$\log(R_t / P_t^Y) = \alpha_4' + (1/\sigma)\log(Y_t / K_t) - [(1-\sigma)/\sigma]\lambda_K \cdot t + \varepsilon_{4,t} \qquad (4')$$

$$\log(W_t / P_t^Y) = \alpha_5' + (1/\sigma)\log(Y_t / L_t) - [(1-\sigma)/\sigma]\lambda_L \cdot t + \varepsilon_{5,t} \qquad (5')$$

$$\log(W_t / R_t) = \alpha_6' + (1/\sigma)\log(K_t / L_t) - [(1-\sigma)/\sigma](\lambda_L - \lambda_K) \cdot t + \varepsilon_{6,t} \qquad (6')$$

As pointed out in Antràs (2004), there are two differences between the specifications in (1') through (6') and those in (1) through (6). First, the aggregate input index F_t is replaced by real output Y_t, and accordingly the price of aggregate input P_t is replaced by the price of output P_t^Y. Second, all six specifications now include a time trend. Therefore, exclusion of the time trend in the regression would result in bias. The only exception is the Cobb-Douglas case (σ =1), where the bias would be zero.

Following Antràs (2004), I use real value added in Canadian business sector to proxy for Y_t, and the corresponding implicit price index to proxy for P_t^Y.

5.3 Estimation Results

Column I in Table 4 presents OLS estimates of equations (1') through (6'). When we compare these estimates with those in column II in Table 1, we could find that they are in general far away from unity, with a wide dispersion around one: from 0.289 to 1.272. Since the standard errors are much larger than those in Table 1, the null hypothesis of a unit elasticity cannot be rejected at the 5% significance level for four of the six specifications, which can be learnt from the t-stats given in the table. Again the Durbin-Watson statistics are very low, which indicates the existence of serial correlation in the residuals.

Table 4 Estimates Allowing for Biased Technological Change

		OLS I	FGLS II	GIV III	AR(2) IV	2SLS V	Saikkonen VI	With Lags VII
1'	σ_1	**0.568**	**0.519**	**0.254**	**0.507**	**0.195**	**0.519**	**0.525**
	S.E.	0.071	0.050	0.250	0.050	0.118	0.121	0.054
	t-stat for $H_0:\sigma_1=1$	-6.072	-9.661	-2.987	-9.782	-6.828	-87.399	-8.741
	R^2	0.789	0.790	0.146	0.962	0.916	0.803	0.948
	D-W	0.275	1.085	1.205	1.836	1.929	0.269	1.069
2'	σ_2	**0.870**	**0.911**	**0.449**	**0.384**	**0.360**	**0.907**	**0.357**
	S.E.	0.095	0.055	0.136	0.134	0.131	0.109	0.111
	t-stat for $H_0:\sigma_2=1$	-1.364	-1.612	-4.041	-4.583	-4.896	-25.341	-5.820
	R^2	0.983	0.944	0.873	0.995	0.995	0.985	0.996
	D-W	0.414	1.334	1.455	1.936	1.934	0.474	1.545
3'	σ_3	**0.289**	**0.507**	**0.232**	**0.304**	**0.196**	**0.062**	**0.379**
	S.E.	0.091	0.065	0.131	0.055	0.127	0.180	0.055
	t-stat for $H_0:\sigma_3=1$	-7.787	-7.531	-5.871	-12.706	-6.313	-67.773	-11.328
	R^2	0.968	0.749	0.405	0.994	0.965	0.966	0.994
	D-W	0.227	1.229	1.025	2.048	0.262	0.279	1.106
4'	σ_4	**0.871**	**0.691**	**0.312**	**0.663**	**0.955**	**1.134**	**0.699**
	S.E.	0.109	0.069	3.678	0.067	0.283	0.295	0.073
	t-stat for $H_0:\sigma_4=1$	-1.185	-4.504	-1.926	-5.016	-0.158	6.107	-4.146
	R^2	0.692	0.762	0.087	0.915	0.675	0.708	0.895
	D-W	0.434	1.150	1.287	1.863	0.467	0.419	1.163
5'	σ_5	**1.224**	**1.026**	**0.867**	**1.517**	**1.113**	**1.200**	**1.418**
	S.E.	0.134	0.061	0.233	0.337	0.130	0.161	0.439
	t-stat for $H_0:\sigma_5=1$	1.673	0.422	-0.762	1.534	0.869	31.986	0.952
	R^2	0.982	0.947	0.875	0.993	0.980	0.980	0.992
	D-W	0.492	1.247	1.254	1.757	0.486	0.476	1.217
6'	σ_6	**1.272**	**0.875**	**1.324**	**0.627**	**1.606**	**-7.483**	**0.625**
	S.E.	0.403	0.111	0.755	0.121	0.950	17.150	0.090
	t-stat for $H_0:\sigma_6=1$	0.675	-1.128	0.245	-3.073	0.638	-5.558	-4.145
	R^2	0.901	0.817	0.571	0.966	0.897	0.934	0.971
	D-W	0.555	1.235	1.221	1.927	0.600	0.797	1.270
No. of Obs.		37	37	36	35	35	34	36

To be consistent with Antràs (2004), column II in Table 4 reports FGLS estimates that apply the Prais-Winston procedure. The results are qualitatively similar to OLS estimates. They are distributed around one in a range from 0.507 to 1.026. For each of the estimates, the FGLS standard error is much lower. However, the null hypothesis of a unit elasticity cannot be rejected at the 5% significance level for three out of the six specifications. In column III, I present estimates based on GIV technique as same as that used in section 4. In contrast to the case in Antràs (2004), the GIV technique seems to make the dispersion of estimates from different specifications much wider, with the lowest estimate being 0.232 and the highest 1.324. And in general the standard errors are very large for GIV estimates.

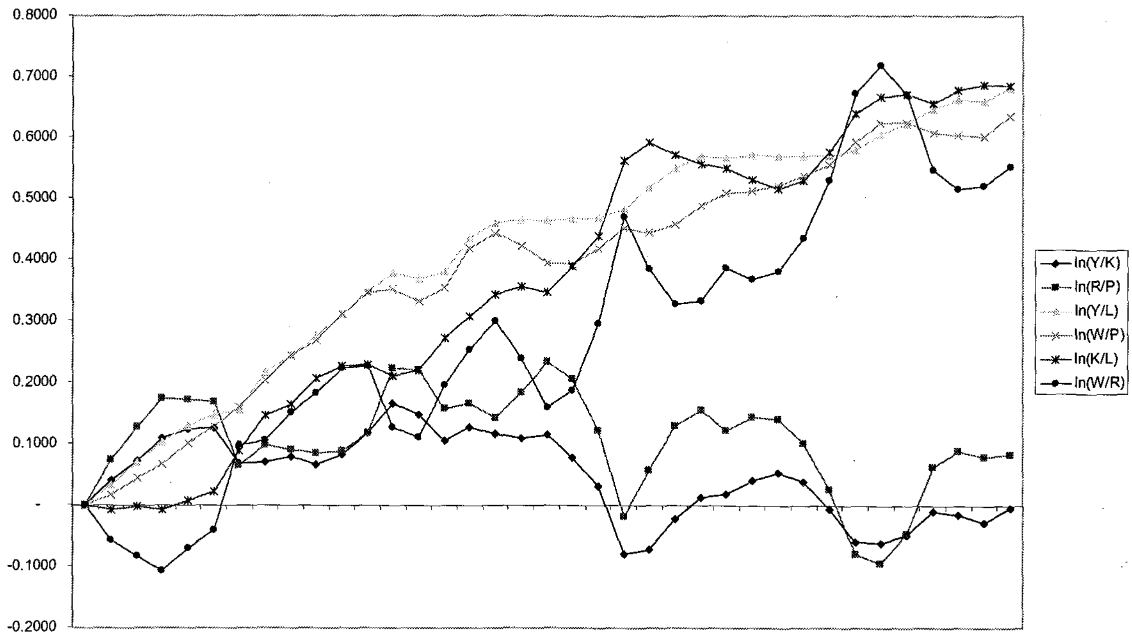
It is likely that the contradiction between the natures of my GIV estimates and those in Antràs (2004) results from the different autocorrelation process of OLS residuals. As mentioned in Antràs, assuming an AR(1) process in autocorrelation structure could leave us as many observations as possible, which is essential for our case of small sample size. On the other hand, he also performed a Ljung-Box test to test whether the residuals in the AR(1) regressions were white noise. He was lucky that his assumption of an AR(1) process was justified by his favorable results of the Ljung-Box tests. However, when I applied the Ljung-Box tests to Canadian data, the null hypothesis of the residuals in specifications (1') through (6') following an AR(1) process was in general rejected at 5% significance level. Especially, it is rejected even at 1% significant level for equations (3') and (5'). In light of this, I imposed an AR(2) process, and ran the regression $\hat{\varepsilon}_t = \rho_1 \hat{\varepsilon}_{t-1} + \rho_2 \hat{\varepsilon}_{t-2} + u_t$, where $\hat{\varepsilon}_t$ is the vector of OLS residuals in Table 4. Ljung-Box tests at up to 5 lags were performed for each of the six specifications leading to no rejections of the null hypothesis of the estimated residuals u_t to be white noise.

Eviews estimates AR models using nonlinear regression techniques. The details can be found in Fair (1984, pp 210 – 214) and Davidson and MacKinnon (1996, pp 314 – 341). Column IV of Table 4 presents the estimates based on AR(2) models. It is apparent that five out of the six estimates are much lower than one. Moreover, they are statistically different from one at 5% significance level.

To deal with the problem of endogeneity accompanied by AR errors, EViews uses a nonlinear least squares procedure as described in Fair (1984). Column V reports the 2SLS estimates. It seems that the method of 2SLS with AR(2) errors does even worse than the GIV technique in bringing estimates together.

The high R-squares and low Durbin-Watson statistics obtained under OLS indicate that the results may suffer from spurious regression bias, just as in section 4. Figure 2 graphs the six series involved in our estimation based on biased technological change, with the logarithms of the variables normalized to zero in 1961.

Figure 2 Nostationarity of the Series



As is similar to Figure 1, we could see that the logarithms of W_t / P_t^Y, Y_t / L_t, W_t / R_t and K_t / L_t all trend upwards, whereas R_t / P_t^Y and Y_t / K_t show a somewhat downward trend. I repeated the unit root tests for the series involved in the estimation of equations (1') through (6'), with the results reported in Table 5. In general the null hypothesis of a unit root cannot be rejected for the level form of the six series, but is rejected for all the six first differenced series. Therefore, we conclude that all the six series are integrated of order one. Then a cointegration test is required to tell whether or not the regressions above are spurious. Part A of Table 6 presents the results of Engle and Granger's (1987) residual-based ADF test. The residuals are obtained from equations (1') through (6'), which include a time trend. And the specification of the test includes zero, one or two lags of the first difference of the residuals. Considering the inclusion of time trend in the equations, I used the "with trend" parameters given in MacKinnon (1991) to compute the critical values, as noted in Antràs (2004). Different from the case in section 4, the null hypothesis of a unit root in residuals cannot be rejected for almost all the six specifications at 5% significance level, with only one exception where the statistic is marginally higher than the 5% critical value. Therefore, it seems that there is no cointegration relationship between the two variables in any specifications. Part B of Table 6 shows the results of Johansen and Juselius (1990) test, which also support our finding of no cointegration.

Table 5 Unit Root Tests

	$\log\left(\frac{Y}{K}\right)$	$\log\left(\frac{R}{P}\right)$	$\log\left(\frac{Y}{L}\right)$	$\log\left(\frac{W}{P}\right)$	$\log\left(\frac{K}{L}\right)$	$\log\left(\frac{W}{R}\right)$	5% Critical Value
ADF 0	-2.799	-2.976	-1.831	-1.679	-1.622	-2.720	-3.539
ADF 1	-3.517	-3.912	-1.980	-2.692	-3.174	-4.841	-3.543
ADF 2	-3.063	-3.244	-1.779	-1.857	-2.475	-3.754	-3.547
PP 2	-2.984	-3.192	-1.834	-1.764	-2.198	-3.110	-3.539
PP 3	-2.953	-3.100	-1.831	-1.654	-2.150	-2.922	-3.539
PP 4	-2.897	-2.997	-1.835	-1.597	-2.017	-2.721	-3.539

	$\Delta\log\left(\frac{Y}{K}\right)$	$\Delta\log\left(\frac{R}{P}\right)$	$\Delta\log\left(\frac{Y}{L}\right)$	$\Delta\log\left(\frac{W}{P}\right)$	$\Delta\log\left(\frac{K}{L}\right)$	$\Delta\log\left(\frac{W}{R}\right)$	5% Critical Value
ADF 1	-4.157	-4.639	-3.845	-5.556	-3.955	-4.995	-2.950
PP 3	-3.864	-4.244	-3.843	-3.387	-3.515	-4.016	-2.947

Table 6 Cointegration Tests

A. Residual-Based Augmented Dickey-Fuller Tests

	Residuals of eq(1')	Residuals of eq(2')	Residuals of eq(3')	Residuals of eq(4')	Residuals of eq(5')	Residuals of eq(6')	5% Critical Value
ADF 0	-1.745	-2.104	-1.139	-2.022	-2.133	-2.379	-3.539
ADF 1	-2.717	-2.817	-2.132	-2.963	-3.486	-3.569	-3.543
ADF 2	-2.272	-2.175	-2.084	-2.286	-2.403	-2.646	-3.547

B. Johansen-Juselius Cointegration Tests

	Max-Lambda				Trace			
Test	r =0 vs r =1		r≤1 vs r =2		r =0 vs r =2		r≤1 vs r =2	
No. of Lags	1	2	1	2	1	2	1	2
log(Y/K) & log(R/P)	14.26	11.68	7.02	5.6	21.28	17.28	7.02	5.6
log(Y/L) & log(W/P)	11.68	12.53	7.98	5.96	19.66	18.49	7.98	5.96
log(K/L) & log(W/R)	19.37	18.8	4.24	5.29	23.6	24.09	4.24	5.29
5% Critical Value	18.96		12.25		25.32		12.25	

Given that there is no cointegration between any pair of variables involved in the estimation, the OLS estimates should be not consistent. However, taking into account of our small sample size and the possible small-sample bias, we cannot make such a firm conclusion. To be consistent with Antràs (2004), I also report the results of applying Saikkonen's (1990) procedure for estimating cointegrating vectors, assuming existence of cointegration relationships. Column IV of Table 4 presents the results of the implementation of Saikkonen's (1990) procedure for $l = 1$ and $p = 1$. And the t-statistics reported are those modified according to the method summarized in Appendix B of Antràs (2004). As noted by Antràs, they should be compared with the associated critical value from a standard normal distribution. The most striking finding from column VI is that I obtained a negative elasticity of substitution from specification (6'), which cannot be interpreted reasonably. Obviously the wrong sign strongly indicates that the assumption of cointegraion does not hold water.

If we believe that there is no cointegration, which seems to be more convincing, then OLS estimates are not reliable because of the spurious regression bias. Following Antràs (2004), I included lagged values of both the dependent and independent variables in the regression as a remedy for the spurious regression bias. As noted by Antràs, this procedure leads to consistent estimates of the elasticity and to t-stats of the hypothesis $\sigma_i=1$ that are asymptotically N(0,1). Column VII of Table 4 presents the estimates resulting from such a procedure. They are largely different from Saikkonen estimates for several specifications, suggesting that the spurious regression biases might be important. The estimates of the elasticity obtained from equations (1'), (2'), (3'), (4') and (6') are substantially lower than one. However, the estimate obtained from equation (5') is much higher than one. As noted in Antràs (2004), the approach of adding lags of both variables in the model is only appropriate under the null hypothesis of no cointegration of the variables. If we check Table 6 again, we could find that the ADF1 statistic for equation (5') is very close to the 5% critical value. And this may be some evidence that there exists cointegration between the two variables in equation (5'). It would be a reasonable explanation especially when we take account of the small size of our sample. Another point worth noting is the similarity between the AR(2) estimates in column IV and the "With Lags" estimates in column VII. It can be regarded as another justification of our imposition of AR(2), instead of AR(1), process for the autocorrelation structure in residuals.

6 Conclusion

For this project, I use Canadian data to replicate Antràs' (2004) paper. That paper illustrates with the U.S. data that ignoring the bias in technological change necessarily leads to an acceptance of the null hypothesis of a unit elasticity of substitution between capital and labor, while the elasticity is likely to be considerably below one with biased technological change. This result is found to be true for Canadian data as well. Under the assumption of Hicks-neutral technological change, my estimation shows that more than half of the estimates of elasticity of substitution between capital and labor are insignificantly different from unity. Allowing for biased technological change, the estimates of elasticity of substitution are in general substantially below one. This implies that a Cobb-Douglas specification of Canada's aggregate production function may be not appropriate.

Appendix

Data Sources: CANSIM Series

L: V715913

W: V715919

K: V715914

R: V715910

F: V715915

P: V715911

Y: V715912

P^Y: V715918

Quality-adjusted L: V2007227

Quality-adjusted K: V2007223

Labor compensation: I724149

Capital compensation: I724145

Population: V1

Nominal wage in non-business sector: V720515

Nominal capital owned by government: V647193

Capital price deflator: V3860236

References

Antràs, Pol (2004), Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution, *Contributions to Macroeconomics*, 4:1.

Arrow, Kenneth J., Hollis B. Chenery, Bagicha Minhas, and Robert M. Solow (1961), Capital-Labor Substitution and Economic Efficiency, *Review of Economics and Statistics*, 43:5, pp. 225-254.

Berndt, Ernst R. (1976), Reconciling Alternative Estimates of the Elasticity of Substitution, *Review of Economics and Statistics*, 58:1, pp.59-68.

Davidson, Russell and James G. MacKinnon (1993), *Estimation and Inference in Econometrics*, Oxford University Press.

Engle, Robert F. and Clive W. J. Granger (1987), Co-integration and Error Correction: Representation, Estimation, and Testing, *Econometrica*, 55:2, pp. 251-276.

Fair, Ray C. (1970), The Estimation of Simultaneous Equation Models With Lagged Endogenous Variables and First Order Serially Correlated Errors, *Econometrica*, 38:3, pp.507-516.

Fair, Ray C. (1984), *Specification, Estimation, and Analysis of Macroecnometric Models*, Cambridge, MA: Harvard University Press.

Gu, W., M. Kaci, J.-P. Maynard, and M. Sillamaa (2003), The Changing Composition of the Canadian Workforce and its Impact on Productivity Growth, in *Productivity Growth in Canada – 2002*, edited by J.R. Baldwin, and T.M. Harchaoui, Ottawa: Statistics Canada Catalogue No. 15-204-XPE.

Hamilton, James D. (1994), *Time Series Analysis*, Princeton, NJ: Princeton University Press.

Harchaoui, T.M, M. Kaci, and J.-P. Mynard (2001), The Statistics Canada Productivity Program: Concepts and Methods, Appendix 1 in *Productivity Growth in Canada*, Ottawa: Statistics Canada Catalogue No. 15-204-XIE.

Harchaoui, T.M. and F. Tankhani (2003), A comprehensive Revision of the Capital Input Methodology for Statistics Canada's Multifactor Productivity Program, in *Productivity Growth in Canada – 2002*, edited by J.R. Baldwin, and T.M. Harchaoui, Ottawa: Statistics Canada Catalogue No. 15-204-XPE.

Johansen, Soren and Katarina Juselius (1990), Maximum Likelihood Estimation and Inference on Cointegration – With Applications to the Demand for Money, *Oxford Bulletin of Economics and Statistics*, 52, pp. 169-210.

Kennedy, Peter (2003), Time Series Econometrics, Simon Fraser University, mimeo.

MacKinnon, J.G. (1991), Critical Values for Cointegration Tests, Chapter 13 in *Long-run Economic Relationships: Readings in Cointegration*, edited by R.F.Engle and C.W.J. Granger, Oxford University Press, pp.267-276.

Saikkonen, Pentti (1991), Asymptotically Efficient Estimation of Cointegration Regressions, *Econometric Theory*, 7, pp.1-21.