ECONOMICS BULLETIN

Is the Size Distribution of Income in Canada a Random Walk?

Baotai Wang
University of Northern British Columbia

Tomson Ogwang
Brock University

Abstract

Several recent studies have investigated whether the size distribution of income in the United States follows a random walk meaning that it has a unit root, with mixed results. In this paper, we investigate the same issue using Canadian national and provincial income inequality data. The investigation is conducted using three different unit root tests. The results suggest modeling the Gini coefficients for Canada and for most provinces as an I(1) is quite reasonable.

The authors wish to thank Ajit Dayanandan, U. L. Gouranga Rao, Danny Cho, and Wimal Rankaduwa for their valuable comments on this study. However, the usual disclaimer applies.

Citation: Wang, Baotai and Tomson Ogwang, (2004) "Is the Size Distribution of Income in Canada a Random Walk?." *Economics Bulletin,* Vol. 3, No. 29 pp. 1−9

Submitted: March 21, 2004. **Accepted:** August 12, 2004.

URL: http://www.economicsbulletin.com/2004/volume3/EB−04C20016A.pdf

1. Introduction

Recently, there has been considerable interest on the issue of whether or not the size distribution of income in the United States follows a random walk over time, or, equivalently, it has a unit root. See, for example, Hayes, Slottje, Porter-Hudak, and Schully (1990), Raj and Slottje (1994), and Rothman (1998), among others. If the size distribution of income follows a random walk, then changes in the size distribution of income are independent, which raises public policy questions about the efficacy of government induced income redistribution schemes.

One reason why many studies have examined whether the size distribution of income in the United States has a unit root is that the United States is one of the few countries with sufficiently long time series data to allow for a meaningful investigation. Recently, reasonably long time series provincial and national data on the size distribution of income in Canada have also become available. Given the ongoing public policy debate on the efficacy of Canada's income redistribution programs, such as income assistance, children's benefits, employment benefits, and old age security, it seems worthwhile to examine whether or not the size distribution of income in Canada follows a random walk. We are not aware of any studies which have investigated this issue in the Canadian case. Since some income redistribution programs differ among the ten Canadian provinces (Alberta, British Columbia, Manitoba, New Brunswick, Newfoundland, Nova Scotia, Ontario, Prince Edward Island, Quebec, and Saskatchewan), it would also be interesting to examine whether the provincial size distribution of incomes follow random walk processes over time.

The purpose of this paper is, therefore, to examine whether the Canadian size distribution of income follows a random walk. To this end, we use annual time series data on the Gini coefficient, as the income inequality measure, for the years 1980-2001. The CANSIM data set used provides time series data on both national and provincial Gini coefficients for the aforementioned period. Given previous evidence of the sensitivity of the unit root tests to the test employed (e.g. Rothman, 1998), we conduct our investigations using three different unit root tests, namely, the ADF test (Dickey and Fuller, 1979, 1981), the PP test (Phillips and Perron, 1998), and the DF-GLS test (Elliot, Rothenberg, and Stock, 1996). Models with and without a time trend are analyzed. Our research follows the tradition of testing for unit roots in macroeconomic time series, which started gaining popularity in the early 1980's.

The rest of the paper is structured as follows: In section 2, we provide brief descriptions of the unit root tests employed. The data are described in section 3. The test results are summarized in the penultimate section and concluding remarks are made in the final section.

2. The unit root tests

The ADF and PP tests for testing the null hypothesis of a unit root have been widely used in empirical work. The DF-GLS test has the advantage of having more power than conventional tests in the sense that it is more likely to reject the null hypothesis of a unit root against a stationary alternative when the alternative is true.

Let $\{w_t\}$, $t = 1,2,...,T$, be a particular time series under consideration, where T is the sample size. The series is integrated of order d (denoted $I(d)$) if it attains stationarity after differencing d times. If the series is $I(1)$ it is deemed to have a unit root or it follows a random walk process. This situation arises if its first difference is $I(0)$.

The ADF test is performed by testing $\delta_0 = 0$ against the one sided alternative $\delta_0 < 0$ in the regression:

$$\Delta w_t = \beta_0 + \beta_1 t + \delta_0 w_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta w_{t-i} + e_t\ , \quad t = 1,2,...,T \tag{1}$$

where e_t is the error term and Δ denotes the first-difference operator. Note that equation (1) incorporates both a constant or intercept (β_0) and a time trend variable t. If the series has a constant term (β_0) but no time trend, the term ($\beta_1 t$) is omitted from equation (1). The optimal lag length (p) can be determined using one of the lag length selection criteria, such as the Schwarz selection criterion.

The PP test is performed by testing $\delta_0 = 0$ against the alternative that the series is stationary in the regression:

$$\Delta w_t = \beta_0 + \beta_1 (t - \frac{1}{2}T) + \delta_0 w_{t-1} + e_t\ , \quad t = 1,2,...,T \tag{2}$$

where the term $(t - \frac{1}{2}T)$ denotes the time trend. If the series has a constant (β_0) but no time trend, the term $\beta_1 (t - \frac{1}{2}T)$ is omitted from equation (2).

The DF-GLS *t*-test is performed by testing $\delta_0 = 0$ in the regression:

$$\Delta w_t^d = \delta_0 w_{t-1}^d + \sum_{i=1}^{p} \gamma_i \Delta w_{t-i}^d + e_t\ , \quad t = 1,2,...,T \tag{3}$$

where w_t^d is the locally detrended series w_t, e_t is the error term, and p is an appropriate number of lags. The locally detrended series w_t^d is computed from

$$w_t^d = w_t - \hat{\beta}_0 - \hat{\beta}_1 t\,, \tag{4}$$

where $(\hat{\beta}_0, \hat{\beta}_1)$ are obtained by regressing $\overline{w}$ on $\overline{z}$, and the variables $\overline{w}$ and $\overline{z}$ are constructed as follows:

$$\overline{w} = [w_1, (1-\overline{\alpha}L)w_2, \ldots (1-\overline{\alpha}L)w_T], \tag{5}$$

$$\overline{z} = [z_1, (1-\overline{\alpha}L)z_2, \ldots (1-\overline{\alpha}L)z_T], \tag{6}$$

where $z_t = (1,t)'$, $\overline{\alpha} = 1 + \frac{\overline{c}}{T}$ and $\overline{c} = -13.5$, and L denotes the lag operator. The local detrending depends on whether we consider a model with a drift (a constant) or, more commonly, a model with both a drift and a time trend. If we consider a model with a drift only, then $\overline{\alpha} = 1 + \frac{\overline{c}}{T}$ where $\overline{c} = -7$ and $w_t^d = w_t - \hat{\beta}_0$. See Elliot, Rothenberg, and Stock (1996) for further discussion.

For all the three tests outlined above, if a particular test accepts the null hypothesis for the series in level but rejects the null hypothesis for the series in first difference, then the series has a unit root, i.e. the series follows a random walk process, which may be a random walk with a constant, or a random walk with a constant and a time trend.

3. The income inequality data

The present study uses national and provincial data on annual Gini coefficients for Canada from 1980 to 2001, collected from Table 2020705 in CANSIM II, the Canadian Socio-economic and Information Management database, compiled by Statistics Canada. This happens to be the longest currently available and most reliable time series data on the Gini coefficient for Canada and for its ten provinces listed above.

The Gini coefficient is a measure of inequality that lies between zero (for perfect equality) and one (for perfect inequality) and whose welfare implications are well known. It is widely used to assess whether re-distributive policies have had an impact. The Gini coefficients used in this study are based on family market income. Market income (or income before taxes) is the sum of earnings from employment and net self-employment income, investment income, and private retirement income. Figure 1 shows plots of the national and provincial Gini coefficients for the years 1980-2001. It can be seen that the Gini coefficients for Canada and for most of its provinces are between 40% and 55%. However, the Gini coefficient for the province of Newfound is much higher than this average range, indicating Newfoundland has most severe income inequality problem in Canada.

The ADF, PP and DF-GLS unit-root tests described above are performed on each provincial or national Gini coefficient series. Even though we believe that one lag is appropriate when annual time series data are used, nevertheless the search for the optimal lag length is conducted over four lags in the case of the ADF test. The DF-GLS test was performed with up to three lags. Using three different unit root tests allows us to test the robustness of the results to the choice of the unit root test and disaggregating the data by provinces allows us to discern any significant differences among provinces.

4. Test results.

Table 1 summarizes the results of the ADF test. The results show that, for the Gini coefficient in level form, the null hypothesis of a unit root is rejected at the 5% and 10% significance levels for New Brunswick and Saskatchewan, respectively, when the test incorporates a constant but does not incorporate a time trend; and at 1% and 5% levels for British Columbia and New Brunswick, respectively, when the test incorporates both a constant and a time trend. If the variable is expressed in first difference form, thereby eliminating the trend, the results show that all series, except those for Manitoba and Nova Scotia, are stationary at the 5% significance level. Overall, the ADF test results show that the Gini coefficients for the provinces of British Columbia, New Brunswick, and Saskatchewan are stationary whereas those for Manitoba and Nova Scotia may be integrated of order 2 or more meaning that the Gini coefficient series may become stationary after differencing at least twice. The Gini coefficients for the provinces of Alberta, Newfoundland, Ontario, Prince Edward Island, and Quebec, and the national Gini coefficient are integrated of order 1 or they follow random walk processes.

The results of the PP test, summarized in Table 2, indicate that for the Gini coefficient in level form, the null hypothesis of a unit root is rejected in favor of the stationary alternative at conventional significance levels in the case of New Brunswick, Newfoundland, Prince Edward Island, and Saskatchewan. If the variables are expressed in first difference form, the null hypothesis of a unit root is rejected at conventional significance levels for all the 10 Canadian provinces and at the national level. These results suggest that the Gini coefficients for New Brunswick, Newfoundland, Prince Edward Island, and Saskatchewan are stationary but the Gini coefficients for the other six provinces (Alberta, British Columbia, Manitoba, Nova Scotia, Ontario and Quebec) and the national Gini coefficient have a unit root or they follow random walk processes.

The results of the DF-GLS test are summarized in Table 3. We first tested each Gini coefficient series under the assumption that the series has a drift only. The test results, reported in the upper portion of the table, show that for the Gini coefficient variable in level form, the null hypothesis of a unit root cannot be rejected for Canada and all its provinces except Alberta, when $p=1$. When the number of lags is increased, all series exhibit non-stationarity in levels. For the Gini coefficient variable in first difference form, the null hypothesis of a unit root is rejected at the 5% and 1% significance levels for all the ten provinces and the national level. Clearly, the test results are quite sensitive to the choice of the number of lags. When the number of lags is increased, the null hypothesis of a unit root cannot be rejected for all the series. As noted above, we believe that one lag is appropriate when annual data are used. Accordingly, it seems reasonable to conclude, on the basis of the DF-GLS test results for $p=1$ under the assumption that each series has a drift only, that the national Gini coefficient and the Gini coefficients for all the provinces, except Alberta, are integrated of order 1 or they follow random walk processes.

We also performed the DF-GLS test on each series under the assumption that the series has a drift as well as a time trend. The test results, reported in the lower part of Table 3, show that in the case of one lag and the Gini coefficient in level form, the null hypothesis of a unit root is rejected for Canada and five of its provinces, namely, Alberta, New Brunswick, Newfoundland, Prince Edward Island, and Quebec. For the Gini coefficient in first difference form, the null hypothesis rejected for all the ten Canadian provinces and the national level. Clearly, the test results in this case are also quite sensitive to the number of lags included in the test.

Table 4 presents a summary of the unit root test results presented in Tables 1 to 3. Two points are apparent from the table. First, the tests are sensitive to the choice of unit root tests and the inclusion of a drift or a time trend, especially in the case of the ADF and DF-GLS tests. The results of the PP test, on the other hand, are robust with respect to the inclusion or exclusion of the trend variable. Second, the results for the Province of Ontario are the most consistent in the sense that all the three unit root tests show that the size distribution of income in Ontario follows a random walk.

Overall, the ADF, PP, and DF-GLS test results suggest modeling the Gini coefficients for Canada and for most provinces (except New Brunswick) as an $I(1)$ series is quite reasonable.

5. Conclusions

In this paper, we investigated whether the size distribution of income in Canada follows a random walk, or, equivalently, has a unit root. The investigation was conducted using three different unit root tests with mixed results. Although the results obtained in this paper may be regarded as preliminary while we await the availability of even longer and better time series data on the Gini coefficient and other measures of the size distribution of income for Canada, nevertheless these results may be regarded as an important first step in addressing such an important topic, which has important public policy implications.

Figure 1. Gini Coefficients for Canada and Its Provinces

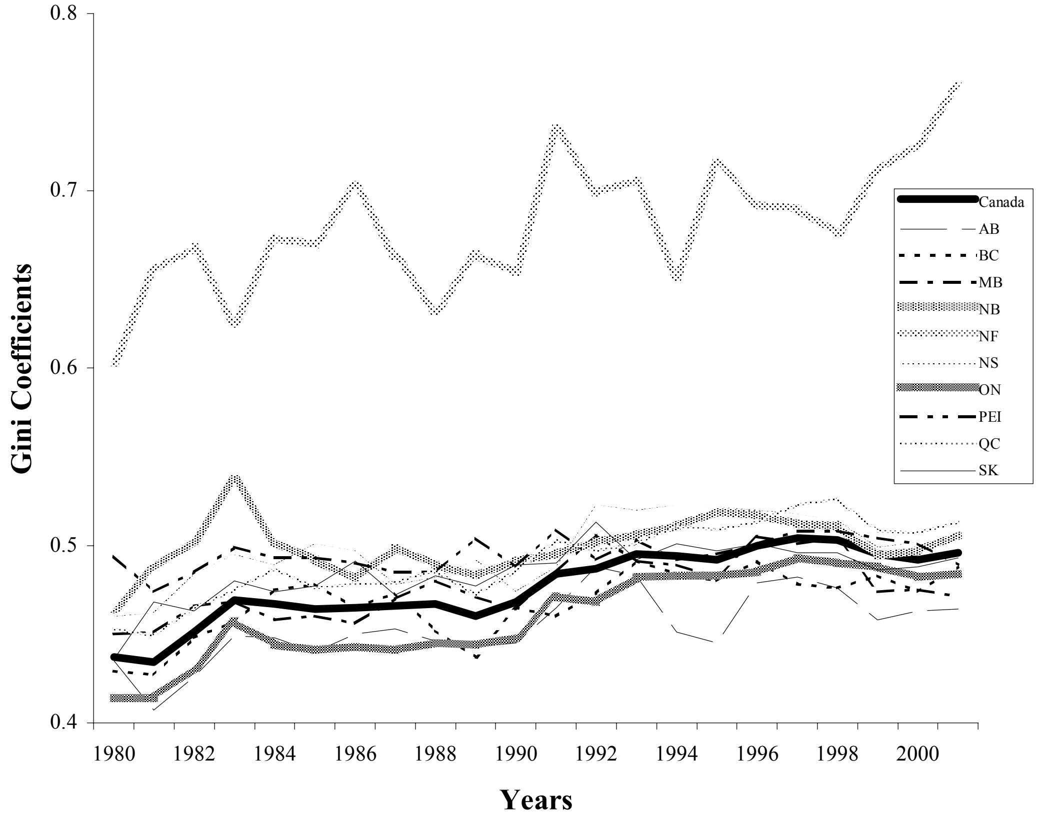

Table 1. Results of the ADF unit root tests

	Constant without trend				Constant with trend			
	Levels		1st differences		Levels		1st Differences	
Canada	-1.8210	[0]	-3.0767**	[2]	-1.9037	[0]	-2.8982	[2]
Alberta	-1.7310	[0]	-4.0574***	[2]	-2.1451	[2]	-4.2305***	[2]
British Columbia	-0.5867	[4]	-4.3954***	[3]	-4.1378***	[4]	-3.8736**	[3]
Manitoba	-1.2024	[3]	-1.8375	[3]	0.1637	[3]	-2.2775	[3]
New Brunswick	-3.3985**	[0]	-3.2840**	[1]	-3.1785*	[0]	-3.0382	[1]
Newfoundland	-1.4282	[1]	-3.7024***	[1]	-2.9299	[1]	-3.6294**	[1]
Nova Scotia	-2.2853	[0]	-1.8293	[2]	-2.0759	[0]	-1.7631	[2]
Ontario	-1.7168	[0]	-3.4052**	[1]	-2.1711	[0]	-3.4629**	[1]
Prince Edward Island	-2.4532	[1]	-3.6320***	[1]	-2.5585	[1]	-3.7958***	[1]
Quebec	-1.6434	[0]	-4.9037***	[0]	-2.6133	[0]	-5.0961***	[0]
Saskatchewan	-2.6803*	[1]	-3.4957***	[1]	-1.8155	[1]	-3.6447**	[1]

Note: The computed t statistics for variables in levels and in first differences are presented in the Table. *, **, and *** indicate significant at the 10%, 5%, and 1% levels respectively. The numbers in the brackets [ ] are the optimal lags, chosen by the Schwarz selection criterion. The critical values for the ADF test are provided by SHAZAM, the econometric software used for performing the test.

Table 2. Results of the PP unit root tests

	Constant without trend		Constant with trend	
	Levels	1st differences	Levels	1st differences
Canada	-1.8060	-3.5973***	-2.0629	-3.8746***
Alberta	-2.1730	-4.3180***	-3.0654	-4.3548***
British Columbia	-2.2564	-5.0111***	-2.6921	-5.0475***
Manitoba	-2.1364	-4.6757***	-2.1104	-4.8276***
New Brunswick	-3.4060**	-4.5538***	-3.2152*	-4.3694***
Newfoundland	-2.8429*	-7.3261***	-4.5079***	-7.1443***
Nova Scotia	-2.2901	-4.2060***	-2.1405	-4.3607***
Ontario	-1.7250	-4.5667***	-2.2101	-4.7874***
Prince Edward Island	-3.3761**	-8.6901***	-4.8856***	-8.8851***
Quebec	-1.6277	-4.9029***	-2.6463	-5.1079***
Saskatchewan	-4.3399***	-9.5045***	-4.4610***	-10.361***

Note: The computed t statistics for variable in levels and in first differences are presented in the Table. *, **, and *** indicate significant at the 10%, 5%, and 1% levels respectively. The critical values for the PP test are provided by SHAZAM, the econometric software used for performing the test.

Table 3. Results of the DF-GLS unit root tests

	Model with a drift					
	$p=1$		$p=2$		$p=3$	
	Levels	1st differences	Levels	1st differences	Levels	1st differences
Canada	-1.517	-3.853***	-0.306	-2.935*	-0.062	-2.034
Alberta	-3.366**	-6.324***	-1.316	-3.931***	-1.050	-2.499
British Columbia	-1.822	-3.362**	-1.238	-1.904	-1.595	-4.354***
Manitoba	-1.786	-3.831***	-1.282	-4.799***	-0.562	-1.821
New Brunswick	-2.216	-3.395**	-2.103	-6.738***	-1.185	-2.017
Newfoundland	-0.976	-3.638**	-0.700	-3.532**	-0.571	-3.310**
Nova Scotia	-1.842	-4.128***	-1.054	-2.019	-1.548	-2.201
Ontario	-0.099	-3.689**	-0.151	-2.096	-0.217	-3.056**
Prince Edward Island	-2.426	-3.689**	-0.263	-2.096	-1.723	-3.056**
Quebec	-1.279	-3.846***	-0.592	-2.568	-0.434	-2.263
Saskatchewan	-0.839	-3.464**	-0.583	-2.935*	-0.397	-2.323
	Model with a drift and a time trend					
	$p=1$		$p=2$		$p=3$	
	Levels	1st differences	Levels	1st differences	Levels	1st differences
Canada	-3.765**	-4.399***	-2.009	-3.122*	-2.244	-2.510
Alberta	-5.800***	-6.465***	-2.398	-4.200***	-1.876	-2.806
British Columbia	-0.739	-2.960*	-0.777	-1.713	-0.303	-3.485**
Manitoba	-2.440	-3.779***	-2.115	-5.007***	-0.989	-1.926
New Brunswick	-2.907*	-3.374**	-3.419**	-6.361***	-3.067*	-2.048
Newfoundland	-2.898*	-3.816***	-3.575**	-4.113***	-2.334	-3.941***
Nova Scotia	-1.973	-4.805***	-1.012	-1.763	-2.475	-2.077
Ontario	-2.662	-3.509**	-2.258	-3.552**	-2.630	-2.158
Prince Edward Island	-2.945*	-3.751**	-3.511**	-2.848	-3.173*	-3.114*
Quebec	-3.275**	-4.460***	-2.253	-3.070*	-2.024	-2.132
Saskatchewan	-1.309	-3.562**	-1.445	-2.890*	-1.322	-2.292

Note: The computed t statistics for variables in levels and in first differences are presented in the Table. The critical values for rejecting the null hypothesis of a unit root are found in Elliott, Rothenberg, and Stock (1996). *, **, and *** indicate significant at the 10%, 5%, and 1% levels respectively. p is the number of lags as defined in equation (3).

Table 4. Summary of the results of the unit root tests

	ADF		PP		DF-GLS (p=1)	
	Constant without trend	Constant with trend	Constant without trend	Constant with trend	Constant without trend	Constant with trend
Canada	**I(1)**	I(0)	**I(1)**	**I(1)**	**I(1)**	I(0)
Alberta	**I(1)**	**I(1)**	**I(1)**	**I(1)**	I(0)	I(0)
British Columbia	**I(1)**	I(0)	**I(1)**	**I(1)**	**I(1)**	**I(1)**
Monitoba	I(>1)	I(>1)	**I(1)**	**I(1)**	**I(1)**	**I(1)**
New Brunswick	I(0)	I(0)	I(0)	I(0)	**I(1)**	I(0)
Newfoundland	**I(1)**	**I(1)**	I(0)	I(0)	**I(1)**	I(0)
Nova Scotia	I(>1)	I(>1)	**I(1)**	**I(1)**	**I(1)**	**I(1)**
Ontario	**I(1)**	**I(1)**	**I(1)**	**I(1)**	**I(1)**	**I(1)**
Prince Edward Island	**I(1)**	**I(1)**	I(0)	I(0)	**I(1)**	I(0)
Quebec	**I(1)**	**I(1)**	**I(1)**	**I(1)**	**I(1)**	I(0)
Saskatchewan	I(0)	**I(1)**	I(0)	I(0)	**I(1)**	**I(1)**

Note: I(0), I(1), and I(>1) denote, respectively, that the series is stationary, that the series has a unit root, and that the series is integrated of order 2 or higher.

References

Elliott, G., T.J. Rothenberg, and J.H. Stock (1996), "Efficient Tests For an Autoregressive Unit Root," Econometrica, 64, 813-836.

Dickey, D. A. and W.A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series With a Unit Root," Journal of the American Statistical Association, 74, 427-431.

__________________________ (1981), "Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root," Econometrica, 49, 1057-1072.

Hayes, K., D.J. Slottje, S. Porter-Hudak and G. Scully (1990), "Is the Size Distribution of Income a Random Walk?" Journal of Econometrics, 43, 213-226.

Phillips, P.C. and P. Perron (1988), "Testing For a Unit Root in Time Series Regression," Biometrika, 75, 335-346.

Raj, B. and D.J. Slottje (1994), "The Trend Behavior of Alternative Income Inequality Measures in the United States From 1947-1990 and the Structural Break," Journal of Business and Economic Statistics, 12, 479-488.

Rothman, P. (1998), "Is the Size Distribution of Income Stationary?" Research on Economic Inequality, 8, 1-12.