Diversification Gains in the Market for Provincial Bonds

Valentina Galvani*
University of Alberta

Aslan Behnamian
University of Alberta

What are the potential benefits of investing in bonds issued by a province for an investor who already holds an efficient portfolio of bonds issued by the remaining nine provinces? Furthermore, how beneficial is the exposure to additional provincial bond markets from the perspective of investors whose core interests reside in maximizing the return to risk bearing? This paper addresses these and similar questions by evaluating the gains from portfolio diversification in the Canadian market for provincial bonds. Our findings indicate that market participants are unlikely to benefit from portfolio diversification. In particular, once realistic trading restrictions are taken into account, the benefits from the exposure to multiple provincial bond markets might be exhausted by investing in the bonds issued by one province alone.

Introduction

This paper quantifies the gains associated with portfolio diversification in the market for bonds issued by the ten Canadian provinces. Diversification benefits from exposure to alternative markets have attracted great attention from individual and institutional investors. Several studies in this field have gauged the advantages of expanding the set of tradable assets from the domestic to the international equity market (among others, Bekaert and Urias, 1996; Errunza et al., 1999; and de Roon et al., 2001). Other works have investigated the gains from portfolio diversification across different classes or sets of assets (e.g., de Roon et al., 2003; Eun, et al., forthcoming). Our analysis contributes to the finance literature on diversification benefits by measuring the gains offered by the same class of assets, namely provincial bonds within the same country. Bonds issued by the Canadian provinces constitute an ideal

*We are grateful to S. Landon and C.E. Smith for generously providing the data for this study, as well as for many helpful conversations. Galvani expresses her gratitude to Tomas K.H. Nilsson for his helpful suggestions. All mistakes are entirely our responsibility

1939-8123/08/1400/00125/$2.50

dataset for an analysis of this type because they are traded in a homogeneous market environment using standardized contracts. To the authors' knowledge, this is the first paper to evaluate the value of portfolio diversification in the market for provincial bonds.

Canadian provinces enjoy significant fiscal independence and can tailor bond issues to their needs. Expenditure and revenue levels vary widely across regional governments, so it might be hypothesized that the differences across provincial economies would be mirrored in their debt markets. If this were the case, provinces should offer differentiated financial products. Consequently, agents could enjoy significant improvements in their investment outlooks by forming portfolios with positions in bonds issued by several, if not by all, provincial governments.

Our results contrast with this view. The following analysis documents the absence of diversification benefits across provincial bond markets for most of the issuing provinces. The evidence is particularly compelling when short-selling restrictions are taken into account. This study indicates that in a restricted trading environment, the advantages stemming from diversification across provinces might vanish for investors holding positions in a benchmark market consisting of one province only. Because anecdotal evidence suggests that shorting provincial bonds is not a feasible trading strategy, we have interpreted the results of our empirical analysis as indicating that individual provinces fail to yield diversification gains with respect to each other.

The average correlation across assets is a rough measure of portfolio diversification benefits: the lower the correlation, the lower the risk associated with a diversified portfolio (e.g., Elton et al., Ch. 4, 2003). In our sample, the average correlation among bonds of different provinces is in the range of 0.9. A standard interpretation of this finding indicates that merging provincial bond markets hardly can be considered an effective strategy for obtaining a substantial reduction in portfolio variance.

Much of the outcome of expanding the set of tradable assets is dependent upon how investors combine their augmented collection of investment possibilities. Therefore, a meaningful comparison of the investment opportunity set before and after the addition of some tradable securities requires a systematic approach that has as its foundation a common portfolio selection procedure. Mean-variance (MV) analysis (Markowitz, 1952) offers a suitable framework for such a comparison.

In an MV framework, diversification gains can be gauged by the extent that the efficient frontier is shifted when the investment opportunity set is expanded. If the MV frontiers of both the benchmark and augmented asset collections are not significantly different, then the benchmark set of investments is said to span the additional investment opportunities. Huberman and Kandel (1987) proposed regression-based tests for spanning.[1]

Tests for spanning can be fruitfully complemented by an assessment of the variations in the risk-bearing compensation that come with an expansion of the investment opportunity set. In the MV framework, the risk compensation of a portfolio is measured by its Sharpe ratio, i.e., the ratio between its excess return and standard deviation. Following the approach proposed by Jobson and Korkie (1982, 1984, 1989), in this study we measure the benefits from portfolio diversification, as perceived by return-to-risk maximizing investors, by the change in the Sharpe ratio that is associated with the exposure to new markets.

The tests for spanning indicate that the gains from portfolio diversification are insignificant. In addition, when trading restrictions are taken into account, the gains that would be obtained from exposure to several provincial bond markets can be exhausted by investing in the debt market for any individual province. An even stronger proposition holds if we start from the perspective of investors primarily interested in maximizing the return to risk bearing. For these investors, adding bonds from additional provinces fails to improve the risk compensation offered by the bonds of other provincial governments, even in an unrestricted trading environment.

Background

Huberman and Kandel (1987) have shown that tests for spanning can be performed by evaluating linear restrictions in a regression of the additional (test) assets' excess returns over the collection of the benchmark excess returns, plus a constant. Excess returns are defined with respect to some risk-free rate, typically identified with the return of some short-term governmental bond. A similar test for spanning has been employed by Bekaert and Urias (1996), with the exception that they substituted net returns for excess returns. Because the present study focuses on the spanning properties of bonds, we prefer, like Bekaert and Urias, to use net returns rather than introduce an arbitrary fixed-income security as a benchmark in our calculations.

Let r_{Nt} and r_{Kt} denote the time-t net returns of N test assets and K benchmark assets, respectively. Testing for spanning is equivalent to evaluating the linear restrictions

$$H_{01}: \alpha = 0_N, \beta 1_K = 1_N, \tag{1}$$

in the multivariate linear model:

$$r_{Nt} = \alpha + \beta r_{Kt} + e_t, \text{ for } t = 1, 2, \ldots, T, \tag{2}$$

---

[1]Successive contributions have linked Huberman and Kandel's testing techniques to tests of the Hansen Jagannathan bounds of the representative investor's marginal rate of substitution. For interested readers, De Roon and Nijam (2001) is an excellent survey of the MV spanning tests.

where:

1_N and 1_K are N and K-dimensional vectors of ones;
0_N is an N-dimensional vector of zeros;
β is an N×K-matrix;
α is an N-dimensional vector,

and the expected value of the disturbances ε_t is zero.

Expressed in prose, there is spanning whenever, up to an orthogonal error term with zero mean, each test-asset return can be written as a portfolio of the benchmark securities whose cost has been normalized to one. Also, it can be shown that there is spanning whenever the test assets fail to improve the compensation to risk bearing and to lower the level of risk obtainable by simply trading the benchmark assets themselves. Tests for spanning summarize two separate measures of diversification gains, briefly reviewed below.

Jobson and Korkie (1982, 1984, 1989) have proven that whenever the linear restrictions

$$H_{02}: \alpha = 0_N, \tag{3}$$

are satisfied in equation (2), then the tangency portfolio of the augmented and benchmark efficient frontiers have the same Sharpe ratio.[2] The Sharpe ratio of the tangency portfolio measures the highest return to risk that can be obtained by forming portfolios with a given collection of risky securities. The Sharpe ratio of the tangency portfolio is henceforth called the *tangency Sharpe ratio*. If H_{02} fails to be rejected, then the test assets fail to improve the risk-bearing compensation offered by the benchmark securities.

It can be shown (e.g., Kan and Zhou, 2001) that whenever the null hypothesis

$$H_{03}: \beta 1_K = 1_N, \tag{4}$$

holds true in equation (2), the global minimum variance (GMV) portfolios of the benchmark and augmented frontiers coincide. The global minimum variance portfolio of a collection of risky assets is defined as the efficient trading position that entails the lowest level of variance and is located to the utmost left of the efficient frontier. As in Van Zijl (1987), we have termed the risk associated with the return of the GMV portfolio *the unavoidable risk*, because no other portfolio of the assigned risky securities yields a return with smaller variance. If the null H_{03} fails to be rejected, then the test assets fail to improve—i.e., to lower—the level of unavoidable risk associated with the benchmark securities.

If the restrictions H_{01} displayed in equation (1) are rejected by the data, then the test assets are not spanned by the benchmark securities. In this case, a separate evaluation of the linear restrictions H_{02} and H_{03}, displayed in equation (3) and equa-

[2]By analogy with the CAPM mis-pricing measure, the components of the vector α are sometimes termed generalized Jensen's alphas.

tion (4), respectively, measures those aspects of the benefits of portfolio diversification that are not captured by the tests for spanning. For example, if H_{02} fails to be rejected but H_{01} is not, then it can be concluded that the test assets fail to improve the benchmark risk-bearing compensation, although the benchmark efficient set is improved significantly by the expansion of the investment possibility set. Similarly, if both the null hypotheses H_{02} and H_{03} are rejected, then the additional securities offer diversification gains in the sense that they improve the efficient risk compensation while lowering the level of unavoidable risk. Of course, if there is spanning, the tangency Sharpe ratio and the GMV return of the augmented and benchmark frontiers will not be significantly different.

Data

This study relies on a dataset containing the end-of-the-month prices and coupons of provincial bonds, issued in Canadian currency, by the ten provinces of Canada: Nova Scotia (NS), Prince Edward Island (PE), Alberta (AB), Manitoba (MN), Newfoundland (NF), Saskatchewan (SK), Ontario (ON), New Brunswick (NB), British Columbia (BC), and the province of Quebec (PQ). Here, the provinces are ordered according to the number of bonds issued, from the smallest to the largest. Only bonds that are completely identified by their maturity date, province of issue, and coupon, are considered.[3]

Our sample covers approximately two decades of monthly observations, from January 1983 to December 2003, and it includes data on 310 provincial bonds. Bonds for which only one return can be calculated were discarded from the sample. The average maturity in the sample is nine years, and the average coupon is 8.9 percent. The time-t holding period net return, expressed by annual rate and in percentage terms, is defined by the following formula:

$$r_t = \left[\left(\frac{p_t - p_{t-1}}{p_{t-1}}\right) \times \frac{365}{N_t} + \frac{C}{P_{t-1}}\right] \tag{5}$$

where:

p_t = The price of a bond with \$100 face value at the end of month t;
N_t = The number of trading days between two price data points in the month under consideration; and
C = The annual coupon.

The overall average return of the provincial bonds in our sample was 9.56. Table 1 provides some summary statistics, including average, robust t-statistic, skewness, and kurtosis of the bond net returns, sorted by issuer.

[3]This unique data library has been used by Landon and Smith (2000, 2007) to examine creditworthiness spillovers across Canadian provinces, as well as to analyze bond returns' seasonal variations (Landon and Smith, 2006).

Table 1—Summary Statistics

Province	Average	t-value	Skewness	Kurtosis	Sharpe I	Sharpe II
NS	8.70	5.64	-0.11	3.64	0.66	1.19
PE	12.03	6.04	-0.37	3.90	0.69	1.30
AB	6.97	4.57	-0.05	4.59	0.84	2.13
MN	8.21	4.83	-0.07	3.60	0.85	1.94
NF	10.8	6.16	-0.26	3.70	0.74	1.59
SK	9.41	5.80	-0.33	4.13	0.88	2.07
ON	8.87	4.50	-0.01	3.86	0.73	1.94
NB	8.98	4.86	-0.19	4.16	0.91	1.70
BC	11.48	5.58	-0.19	3.71	0.82	1.89
PQ	10.14	5.40	-0.10	4.13	0.93	2.30

The robust t-statistics indicate that average returns are significantly different than zero in each province, with a 5 percent significance level. The sample's kurtosis shows that our time series displays rather thin tails, and average skewness suggests that returns are fairly symmetrically distributed around the mean. The Jaque-Bera test (unreported), however, shows that, in each province, less than 40 percent of the bond returns are normally distributed.

The last column of the above table reports the tangency Sharpe ratio, calculated with respect to a risk-free asset with a net rate of return equal to zero.[4] Differences in the tangency Sharpe ratios depend on the position of the risk-free rate relative to the GMV return, as well as on the curvature of the efficient frontier. Averaging over all provinces, the tangency Sharpe ratio is 0.8. In words, for each extra unit of risk as measured by standard deviation, there is a corresponding average increase of 0.8 in the net annual return. For comparison, the Sharpe ratio associated with the monthly returns of the S&P/TSX composite over the time interval covered by our sample is 0.18. Expressed in prose, each extra unit of risk is rewarded with the increase of 0.18 in annual net return.

Bonds have a finite life defined by the contractual maturity date, which implies that most of the securities we have considered drop from our sample at a certain point. One way of handling finitely-lived securities in a study such as ours is to set the asset returns after maturity to zero, an approach utilized in a related study (Behnamian, 2006). This approach has been subsequently abandoned, however, due to concerns regarding possible biases caused by sample unbalance.[5]

An alternative method for handling samples of finitely-lived securities is to focus on long-term contracts, as expiration is an infrequent occurrence for this class of assets. In our sample, however, the long-term contracts constitute a poor proxy of the whole provincial bond market because they are relatively sparse with respect to the total collection of bonds available at each price point. Also, we are not inclined to

[4]If the risk-free net return is zero, then excess and net returns coincide. Hence, considering net returns in this case is equivalent to utilizing excess returns calculated with respect to a risk-free asset bearing a zero interest rate.

[5]We thank S. Landon for raising this point.

discard the medium- and short-term bonds from our sample, because such a sample selection might distort our analysis with respect to the existence of diversification benefits. This could cause a bias because certain trades that provide diversification gains might require the combination of a variety of bond maturities in order to be executed.

Following the methodology of previous contributions (Bessembinder, 1992; de Roon et al., 2000), we have handled this problem by calculating the continuous time series of monthly returns for the two nearest-to-maturity contracts. These time series of returns, two per province, are defined on a moving source: a position is taken in the nearest-to-maturity bond until its maturity month, at which time the position changes to the following nearest-to-maturity contract. An analogous procedure then defines the time series of returns for this subsequent nearest-to-maturity bond.

Another way of handling finitely-lived contracts is to construct indexes.[6] We aggregate bonds from each province into a provincial, equally weighted (EW) portfolio with monthly rebalancing. A reason for looking at such portfolios is that their performance is supposed to measure the return from investing in the average portfolio constituent. By extension, the return from the average portfolio constituent can gauge the average return of the aggregate market.

The average correlation between the provincial moving-source time series of returns and the individual provincial bonds was between 0.7 and 0.8. By means of comparison, the average correlation between the EW indexes and their constituents is never less than 0.9.[7] We therefore are inclined to conclude that each provincial bond market is represented more suitably by an EW index than by the moving-source time series of returns.

The average correlation between the EW indexes and their constituents is reported in the first column of Table 2, while the second column lists the EW indexes' average returns.

Table 2—EW Indexes

Province	Average Correlation	Average
NS	0.93	9.94
PE	0.93	10.86
AB	0.94	8.59
MN	0.90	9.21
NF	0.92	11.01
SK	0.91	9.37
ON	0.93	8.72
NB	0.94	9.63
BC	0.93	9.68
PQ	0.92	9.83

[6]In their study of diversification benefits in emerging bond markets, Chiang et al. (2007) utilize the J.P. Morgan Emerging Market Bond Index.

[7]Of course, the calculation of each correlation coefficient is restricted to the time period in which the bond return is available.

Robust t-statistics indicate that the average return of the EW index for each province is significantly different than zero. Skewness and kurtosis (unreported) suggest that these indexes are also rather symmetric around the mean, with relatively thin tails.

Individual Provinces

This section presents the results of our testing of diversification gains. We test the null of spanning (H_{01}) and of the equality of the tangency Sharpe ratio (H_{02}). An evaluation of the changes in the GMV return (H_{02}) is included only when it clarifies the results of the previous two tests.

The linear restrictions corresponding to the null hypotheses H_{0i} for i = 1, 2, 3 are evaluated with asymptotic and finite sample tests. The asymptotic tests utilized for this purpose are the Lawley-Hotelling trace statistic (a multivariate Wald test), the Wilks' Λ-criterion, and the Pillai trace statistic (e.g., Berndt and Savin, 1977). The variance in the parameter estimates that were used to generate these statistics is corrected for heteroskedasticity and autocorrelation.[8] The resulting statistics are compared with their asymptotic and empirical distributions, and the empirical distribution is estimated using a semi-parametric bootstrap (Davidson and MacKinnon, Section 4.6, 2004). In addition, finite-sample tests of the same linear restrictions (Jobson and Korkie, 1989) also are calculated. In view of the fact that these testing techniques all lead us to the same conclusion, we report only the asymptotic P-value of the Wald test.[9]

Table 3 Panel A reports the Wald test P-value for the tests of diversification benefits summarized by the null hypotheses H_{01}, H_{02}, and H_{03}. The linear model considered matches each province against the other nine. These tests gauge the extent to which individual provincial markets are heterogeneous with respect to the other nine provinces taken in aggregate.

Table 3 Panel B reports the change in the tangency Sharpe ratio (ΔShR) and in the standard deviation of the GMV return (ΔGMV) obtained by adding each province to the other nine ones. The Sharpe ratio is calculated with respect to a zero-rate, risk-free asset. By design, the expansion of the investment possibilities set increases the Sharpe ratio and lowers the level of unavoidable risk. This explains why the statistics listed in Table 3 Panel B are of the opposite sign.

To put these values into context, the average Sharpe ratio of the frontier generated by nine EW indexes is 0.59, and the average GMV standard deviation is 11.4. Therefore, the values in Table 3 Panel B indicate that adding an individual province to the other nine entails changes in the values of the tangency Sharpe ratio and of the

---

[8]The autocorrelation matrices of the Newey-West (1987) variance estimator were calculated assuming a lag of $T^{1/3}$ months, where T is the sample size.

[9]See also Stewart (1995).

GMV standard deviation that are of the order of 1 percent and 3 percent, respectively.[10]

Table 3—Tests of Diversification Gains

Panel A: Wald Test P-value for the Tests of Diversification Benefits

Province	H_{01}	H_{02}	H_{03}
NS	0.02	0.10	0.00
PE	0.19	0.82	0.17
AB	0.00	0.01	0.00
MN	0.00	0.10	0.00
NF	0.17	0.49	0.06
SK	0.05	0.15	0.01
ON	0.06	0.52	0.02
NB	0.47	0.25	0.24
BC	0.87	0.74	0.86
PQ	0.00	0.25	0.00

Panel B: Change in the Tangency Sharpe Ratio (ΔShR) and in the Standard Deviation of the GMV Return (ΔGMV) Obtained by Adding Each Province to the Other Nine

Province	ΔShR	ΔGMV
NS	0.00	-0.28
PE	0.00	-0.10
AB	0.07	-0.84
MN	0.01	-0.83
NF	0.00	-0.17
SK	0.00	-0.25
ON	0.01	-0.64
NB	0.00	-0.12
BC	0.01	0.00
PQ	0.00	-0.51

Panel C: Average Correlation between Each Provincial EW Index and the Remaining Nine Provincial EW Indexes and Number of Pairwise Comparisons in which the Null of Equality of the EW Indexes' Return Distributions Fails to Meet the Significance Level of 5 Percent

Province	Average Correlation	Kolmogrov-Smirnov (KS)
NS	0.94	2
PE	0.92	2
AB	0.89	8
MN	0.90	1
NF	0.92	2
SK	0.94	1
ON	0.84	4
NB	0.90	1
BC	0.93	1
PQ	0.92	2

[10]For the sake of simplicity, Table 3 Panel B does not list the percentage changes in the GMV portfolio mean. Under hypothesis H_{03}, however, both the mean and the standard deviation of the GMV return remain unchanged when the set of traded assets is expanded. In contrast, a significant change in the mean of the GMV portfolio might suffice to reject H_{03}.

Table 3(cont.)—Tests of Diversification Gains

Panel D: Percentage of Comparisons (rounded to the closest integer) in which the Null Hypothesis of Spanning Failed to be Rejected for a Significance Level of 0.05. Frequencies (in percentage terms) are Calculated According to a Baseline Model, in Which the Number of Provinces Aggregated Together into a Given Benchmark Market K ranges from 1 to 9

Province	H_{01}	H_{02}	H_{03}
NS	42	91	36
PE	50	100	61
AB	5	17	2
MN	32	80	24
NF	37	100	31
SK	21	50	15
ON	45	100	25
NB	72	87	66
BC	73	100	71
PQ	28	94	27

Panel E: Frequencies Calculated for K Ranging from 1 to 4

Province	H_{01}	H_{02}	H_{03}
NS	45	97	40
PE	44	99	60
AB	3	23	1
MN	42	84	34
NF	35	99	33
SK	18	37	14
ON	44	100	27
NB	60	82	58
BC	64	99	60
PQ	32	95	33

Panel F: Frequencies Calculated for K Ranging from 5 to 9

Province	H_{01}	H_{02}	H_{03}
NS	39	82	32
PE	56	100	62
AB	6	10	3
MN	21	76	15
NF	40	100	28
SK	25	63	15
ON	45	100	22
NB	82	91	75
BC	81	100	83
PQ	23	94	21

At a 5 percent significance level, the values in Table 3 Panel A indicate that, with the exception of AB, the addition of the test market fails to improve the risk-bearing compensation beyond that of an efficient, optimal portfolio of the remaining nine provinces. In six of ten provinces, this lack of diversification benefits is confirmed by the tests for spanning. In the remaining four provinces, the rejection of the spanning hypotheses is driven by a significant change in the GMV return.

The analysis of the results reported in Tables 3_a and 3_b suggests that tests of diversification gains are sensitive to small changes in the GMV returns. By means of example, consider the expansion of the investment possibility set associated with adding the province of NS to the other nine provincial bond markets. Our calculations indicate that a reduction of the level of unavoidable risk by about a quarter of a standard deviation (0.28) of the annualized returns decreases the overall level of risk, from the benchmark value of 11.37 standard deviations to 11.09. While this change bears little or no economic significance, the associated statistical test reported in Table 3 Panel A indicates that NS should offer significant diversification gains in terms of decreasing the level of unavoidable risk.

Kan and Zhou (2001) have analyzed the relative performance of the tests of the null hypotheses H_{02} and H_{03}. Relying on simulations, they argue that changes in the GMV return that are of negligible importance from a sheer economic perspective tend to drive a rejection of the null of spanning. In this sense, tests for spanning are prone to overestimating diversification gains.[11]

In view of Kan and Zhou's analysis, we take the P-values of Table 3 Panel A as supporting the conjecture that individual provinces fail to offer significant diversification benefits over the aggregate market where gains are measured by the shifts of the efficient frontier.

Standard portfolio management theory states that investors can improve portfolio performance by diversifying across different assets. The argument is that agents can decrease their level of risk by allocating funds to imperfectly correlated securities. Accordingly, the correlation between the assets held in the benchmark portfolio and the constituents of the test market is a rough measure of diversification gains.

The first column of Table 3 Panel C reports the average correlation between each provincial EW index and the remaining nine provincial EW indexes. The majority of these values scored highly, in the range of 0.9, which indicates that individual bond markets are highly correlated among one another. Hence, our interpretation of the empirical evidence reported in Table 3 Panel A is in line with a traditional portfolio management assessment of the diversification benefits in the provincial bond market.

From a general perspective, provincial bonds should be considered differentiated financial products whenever their underlying return-generating processes are significantly different. In order to assess whether the bond return distributions vary between provinces, we use the Kolmogorov-Smirnov (KS) test to compare the returns of the EW indexes.

The column KS in Table 3 Panel C lists the number of pairwise comparisons, of a total of nine, in which the null of equality of the EW indexes' return distributions fails to meet the significance level of 5 percent. These values indicate that the returns

[11]To obviate this shortcoming, Kan and Zhou suggested adjusting the acceptance levels of the tests of the null hypotheses H_{02} and H_{03} against the rejection of the null of the GMV returns' equality. In this study, we maintain the 5 percent significance level for both tests.

of the provincial indexes, with the exception of AB, seem to be associated with comparable data-generating processes.[12]

To this point, our analysis has concerned diversification gains offered by an individual province when it is added to the remaining nine issuers in aggregate. Hereafter, we will refine our analysis by matching each province to subgroups containing the same collection of benchmark assets only. The purpose of this comparison is to ascertain whether the gains offered by each individual province can be obtained by trading bonds issued by fewer than nine other provinces.

Before proceeding, a small note of caution is in order. Let us assume that a provincial market is spanned by a collection of benchmark markets. For example, let PE be spanned by both BC and PQ, separately. It is tempting to aggregate the benchmark markets and conclude that BC merged with PQ spans PE. At a given significant level, this need not be the case, because the spanning tests for the merged market is a joint test of the hypotheses of the pairwise comparisons. For example, the P-values of the Wald test of H_{01}, when PE is the test market and BC and PQ are the benchmark assets, are 0.012 and 0.02 respectively. The corresponding P-value for the merged BC and PQ markets is 0.00. Thus, at the significance level of 1 percent, PE is spanned by BC and PQ separately, but not by the merged market. Similarly, a collection of benchmark assets that spans two test markets need not span the merged test markets. Put differently, tests for spanning are not robust to aggregation. A similar proposition holds for the other metrics of diversification gains introduced in earlier.

A full assessment of the diversification benefits entailed by each individual province requires the evaluation of the null hypotheses H_{0i} for $i = 1, 2, 3$ for all levels of aggregation of the remaining nine provinces. Consequently, the format in which the results of these tests are presented must be different from the format used in Table 3 Panel A, due to the fact that there are approximately 500 combinations to be considered for each province.

Table 3 Panel D reports the percentage of comparisons (rounded to the closest integer) in which the null hypothesis of spanning failed to be rejected for a significance level of 0.05.[13] The frequencies (in percentage terms) in Table 3 Panel D are calculated according to a baseline model, in which the number of provinces aggregated together into a given benchmark market K ranges from 1 to 9. The frequencies in Table 3 Panel E and Table 3 Panel F are calculated for K ranging from 1 to 4 and from 5 to 9, respectively.

---

[12]The KS test can also be used to evaluate the first-degree stochastic dominance relationship. It may be of some interest to note that in our sample, the EW index of AB dominates with first-degree stochastic dominance the return distribution of the other EW indexes in eight of nine pairwise comparisons.

[13]The full results of applying each benchmark market combination to each test province are available from the authors upon request.

These frequencies indicate that expanding the investment opportunity set with any additional province fails to improve the return to risk-bearing offered by the benchmark market. Significant exceptions to this overall assessment are AB and, to a lesser degree, SK. Tables 3_e and 3_f indicate that this result is robust with respect to the size of the benchmark market. In addition, insignificant changes to the tangency Sharpe ratio are pervasive throughout the pairwise comparisons between provinces (unreported). In view of this finding, we have concluded that individual provinces do not offer reciprocal gains in terms of risk compensation.

Once we turn our attention to the frequencies associated with the tests for spanning (H_{01}), our results are less univocal. For seven of ten provinces, the data suggest that there are diversification benefits in a substantial proportion of comparisons, in many cases close to 50 percent. This holds true across varying sizes of the benchmark market (Tables 3_e and 3_f). Exceptions are the test markets NB and BC, which are spanned in more than 50 percent of assessments, a result mainly driven by cases in which the benchmark market aggregates at least five provinces.

As mentioned before, the tests for spanning are somewhat prone to overestimating the significance of diversification benefits offered by an expansion of the investment opportunity set. Bearing this observation in mind, our analysis seems to indicate that it is possible to exhaust the gains from cross-province diversification by investing in bonds issued by a handful of provinces. Notable exceptions, however, are the bonds issued by the province of Alberta. Regardless of the benchmark market size, the addition of Alberta bonds seemed to improve both the risk-bearing compensation and the level of unavoidable risk attainable by trading efficient portfolios of bonds issued by the other provinces only.

Short-Selling Restrictions

In this section, we will evaluate the diversification gains offered by each Canadian province's bonds if we assume that short selling is not allowed. Short selling is an investment strategy involving the sale of an asset that is not owned by the investor. It is a trade that is possible for some—but not all—investment assets, in which market makers lend the asset in question to the investor that wants to sell it short. Futures contracts on bonds are thought to facilitate taking short positions when market makers fail to support short sales. The absence of a derivative market for bonds issued by Canadian provinces suggests that selling short these debt securities might be unfeasible. Therefore, the assessment of the value of portfolio diversification for provincial bonds should take into account short selling restrictions.

A constrained market environment might dilute potential gains arising from portfolio diversification. This is because trade limitations might make some portfolios that would otherwise reap the advantages of an expanded investment opportunity set unavailable for trade. On the other hand, however, the very same constraints might make certain efficient trades involving the benchmark assets unavailable as well. In this case, exposure to the test market might allow for the formation of portfolios with superior performance given the restrictions imposed by the market environment.

The measures of diversification gains reviewed earlier make the assumption that agents can short sell both the test and benchmark assets. Recently, de Roon et al. (2001) have shown that the tests for spanning can be modified to account for short-selling restriction: the linear restrictions to spanning in equation (1) are substituted by a collection of inequalities of the form $v\alpha - \beta 1_K \leq 1_N$ for each zero-beta return v, where α and β are the coefficients of the linear model in equation (2).[14] When short-selling restrictions apply to both the benchmark and test markets, the parameters α and β in these inequalities are substituted with the corresponding estimates $\alpha^{(v)}$ and $\beta^{(v)}$, obtained from a modified version of the baseline model. This modified model consists of a linear regression equivalent to the one in equation (2), in which the benchmark assets are restricted to those constituting the efficient portfolio whose zero-beta rate is v, under the no-short-sales assumption. Identifying these assets is a matter of solving a quadratic optimization problem with inequality constraints. The Elton, Gruber, and Padberg algorithm for the identification of the tangency portfolio under no-short-sales restriction (e.g., Elton et al., Ch. 9, 2003) indicates that the relevant range of the zero-beta return v can be partitioned into intervals in which $\alpha^{(v)}$ and $\beta^{(v)}$ remain unchanged. Finally, testing for spanning is equivalent to jointly evaluating the inequalities $v\alpha - \beta 1_K \leq 1_N$ at the extremes of these intervals of v. For further details, see de Roon et al. (2001) and Kodde and Palm (1986).

Table 4 reports the results of the tests for spanning under short-selling restrictions. For each province, the P-value below is obtained from the empirical distribution of the test statistic estimated with a semi-parametric bootstrap (Davidson and MacKinnon, Section 4.6, 2004).

Table 4—Tests for Spanning under Short-Selling Restrictions

Province	P-Value
NS	0.34
PE	0.46
AB	0.17
MN	0.96
NF	0.42
SK	0.67
ON	0.42
NB	0.51
BC	0.64
PQ	0.24
Average	0.48

[14]In the standard deviation-expected return plane, the zero-beta rate v of an efficient return r is the intercept of the line tangent to the efficient frontier in *r* (e.g., Cochrane, Ch. 6, 2005).

The P-values listed above univocally indicate that once trading restrictions are taken into account, each province is spanned by the aggregate market of the remaining nine provinces.

To complete our analysis, we evaluate the tests for spanning under short-selling restrictions, with a benchmark market consisting of only one province. There are nine pairwise comparisons per province. The decision to reject the null of spanning was based on the 5 percent critical value of the empirical distribution of the tests. The results of this analysis can be summarized in a single sentence: each province is spanned by any of the other nine provinces in a pairwise comparison. Put differently, for each province, the rate at which the null of spanning fails to be rejected is 100 percent when considering pairwise comparisons between provinces. We take these results as compelling evidence of the lack of diversification gains across provincial bond markets.

Aggregate Markets

The preceding sections have investigated the diversification gains associated with an extension of the benchmark market by an individual province. Because the measures of diversification benefits under consideration to this point are not robust when test assets are aggregated, the results thus far stop short of suggesting that aggregated provincial bond markets are unlikely to offer diversification gains. In order to complete the picture, we evaluate diversification gains associated with exposure to the market for bonds issued by multiple provinces.

A word of caution is in order when selecting the size of the aggregated test and benchmark markets. Kan and Zhou (1999, Table I) have shown that the test assets should not outnumber the benchmark investment possibilities, as such a scenario would create the possibility of over rejecting a true null. Hence, our assessment is limited to all partitions of the ten provinces in which there are at least as many benchmark securities as test assets.

Each row of Table 5 shows the frequency (in percentage) that the null hypotheses H_{0i} for i=1, 2, 3 fail to be rejected when we examine all comparisons between a test market of N indexes and a benchmark market constituted by 10 - N indexes.

Table 5—Comparisons between a Test Market of N Indexes and a Benchmark Market Constituted by 10 - N Indexes

N	H_{01}	H_{02}	H_{03}
2	27	100	13
3	15	97	5
4	6	96	1
5	1	92	0

Table 5 suggests that adding more than two provinces to the benchmark market does create significant diversification benefits, more specifically benefits concentrated in a decrease of the overall level of unavoidable risk. The frequencies relative to the null hypothesis H_{02} indicate that in the overwhelming majority of comparisons, adding additional provinces fails to enhance the benchmark return to risk. That is, aggregated markets of provincial bonds fail to offer diversification benefits when gains are measured according to the return to risk-bearing.

The percentages listed reveal that diversification benefits increase with the size of the test market. The frequency with which the null hypotheses of spanning were rejected confirms our initial intuition that gains from portfolio diversification are less likely to be significant when adding a small number of investment opportunities, as opposed to adding a large number.

Conclusions

This study has evaluated the diversification benefits in the market for bonds issued by the Canadian provinces. The empirical analysis presented here relies on the MV tests for spanning to measure gains achieved through portfolio diversification. Other metrics considered here gauge the changes in the level of risk compensation and in the level of unavoidable risk associated with expanding the investment possibility set.

We have shown that it is possible to exhaust the gains from cross-province diversification by investing in bonds issued by a handful of provinces. Our analysis yields an even stronger conclusion when short-selling restrictions are taken into account. In a restricted trading environment, the exposure to several provincial bond markets need not entail significant benefits over investment strategies limited to bonds issued by an individual province. In this sense, our work indicates that the average performance of a provincial bond market can be replicated by bonds issued by any other provincial government. Furthermore, the assessment of the risk-bearing compensation associated with optimal portfolios indicates that agents whose core interests reside in achieving a high return to risk are unlikely to benefit from the trade of bonds issued by more than one province.

The tests for spanning used in this study were derived from the framework of the static MV portfolio selection procedure. Hence, our results compare the investment opportunities offered by buy-and-hold portfolios only. Nevertheless, we would argue that our assessment of the gains from diversification is robust to the introduction of dynamic investment strategies.

Clearly, there is no way to account for all of the dynamic trading strategies available to market agents. However, investors can be thought to allocate their wealth in dynamic portfolios that incorporate the relevant pricing information as summarized by a few measurable indexes. These portfolios are obtained through the interaction of the selected information variables and the returns of the available assets in order to obtain the so-called scaled returns of managed portfolios (see also Cochrane, Ch.8, 2005). Once these trading strategies are added to the pool of assets available for trade, the static MV portfolio selection procedure yields dynamic portfolios that are efficient in the sense defined by Markowitz (1952).

The MV tests of diversification benefits also can be modified to accommodate scaled returns (Bekaert and Urias, 1996). These testing techniques depend on the realization of information variables that account for conditioned information and are therefore named *conditional spanning tests*.[15]

This paper's assessment of the diversification gains in the market for provincial bonds was conducted within the unconditional MV framework. Our results suggest that the benefits of diversification across provincial bond markets are unconditionally insignificant. An inspection of the moment conditions implying the conditional spanning tests introduced in Bekaert and Urias (1996) and by de Roon et al. (2001) reveals that unconditional spanning implies spanning within the conditional framework. This is true whether short-selling restrictions are imposed or not. Hence, our findings are robust with respect to the introduction of dynamic trading strategies that account for changes in the market information structure.

To the extent that trading prices, and thus returns, mirror the collective assessment of the value of an asset, the absence of diversification gains across provinces indicates that the market places similar values on bonds issued by all provincial governments. Therefore, the results presented in this paper suggest that the provincial bond markets offer homogeneous financial products. This lack of heterogeneity goes beyond the high cross-province correlation of bond returns, and in fact, our analysis suggests that individual provincial bond markets fail to offer significant gains with respect to one another, even when assets are combined using optimal investment strategies.

Trends in the overall national economy or in federal monetary policies are likely to affect the provincial debt markets in a uniform fashion.[16] One therefore might conjecture that differences in the rate of return offered by the provincial bond markets are to be attributed to issuers' idiosyncratic default risks, as perceived by market participants. From this perspective, the results of this study tell a story of ex-post homogeneity of the perceived default risk of provinces.

This view is not supported by the wide range of credit ratings attributed to provincial bonds by popular credit-rating agencies. For example, in the vast majority of the months in our sample, the ranks of the ten provinces were spread over seven of the eight Moody's credit-rating categories. (See also Landon and Smith, 2000.) In view of the empirical evidence reported in this paper, a promising direction of further research would be to quantify the extent to which issuers' credit ratings are mirrored by ex-post returns.

---

[15]Other conditional spanning tests are reviewed in de Roon and Nijman (2001).

[16]Because the bonds we have examined are issued in the same currency, returns are not affected by exchange rate fluctuations, which might introduce an exogenous source of return differentiation.

References

1. Alexander, G.J., “Short Selling and Efficient Sets,” *Journal of Finance*, 48, no. 4 (1993), pp. 1497-1506.

2. Bessembinder, H., “Systematic Risk, Hedging Pressure, and Risk Premiums in Futures Markets,” *The Review of Financial Studies*, 5(41992), 637-667.

3. Behnamian, A., “Diversification Benefits in the Canadian Market for Provincial Bonds,” ECON 900 student paper, University of Alberta (unpublished).

4. Bekaert, G., and M.S. Urias, “Diversification, Integration and Emerging Market Closed-end Funds,” *Journal of Finance*, 51, no. 3 (1996), pp. 835-869.

5. Berndt, E.R., and N.E. Savin, “Conflict among Criteria for Testing Hypotheses in the Multivariate Linear Regression Model,” *Econometrica*, 45, no. 5 (1977), pp. 1263-1277.

6. Chiang, K.C.H., C.H. Wisen, and T. Zhou, “Emerging Market Bonds as an Asset Class: Mean-Variance Spanning,” *Journal of Investing*, 16, no. 3 (2007), pp. 104-110.

7. Cochrane J.H., *Asset Pricing*, 2nd ed. (Princeton, NJ: Princeton University Press, 2005).

8. Davidson, R., and J.G. MacKinnon, *Econometric Theory and Methods* (Oxford University Press, 2004).

9. De Roon, F.A., and T.E. Nijman, “Testing for Mean-Variance Spanning: A Survey,” *Journal of Empirical Finance*, 8, no. 2 (2001), pp. 111-156.

10. De Roon, F.A., T.E. Nijman, and C.H. Veld, “Hedging Pressure Effects in Futures Markets,” *Journal of Finance*, 55, no. 3 (2000), pp. 1437-1456.

11. De Roon, F.A., T.E. Nijman, and B.J.M. Werker, “Testing for mean- Variance Spanning with Short Sales Constraints and Transaction Costs: The Case of Emerging Markets,” *Journal of Finance*, 56, no. 2 (2001). pp. 721-742.

12. De Roon, F.A., T.E. Nijman, and B.J.M. Werker, “Currency Hedging for International Stock Portfolios,” *Journal of Banking & Finance*, 27 (2003), pp. 327-349.

13. Elton E.J., M.J. Gruber, S.J. Brown, and W.N. Goetzmann, *Modern Portfolio Theory and Investment Analysis*, 6th ed. (Hoboken, NJ: Wiley and Sons, 2003).

14. Errunza, V., K. Hogan, and M.V, Hung, “Can the Gains from International Diversification be Achieved without Trading Abroad?” *Journal of Finance*, 54, no. 6 (1999), pp. 2075-2107.

15. Eun, C.S., W. Huang, and S. Lai, “International Diversification with Large- and Small-Cap Stocks,” *Journal of Financial and Quantitative Analysis* (forthcoming).

16. Huberman, G., and S. Kandel, “Mean-Variance Spanning,” *Journal of Finance*, 42, no. 4 (1987), pp. 873-888.

17. Jobson, J.D., and B. Korkie, “Potential Performance and Tests of Portfolio Efficiency,” *Journal of Financial Economics*, 10, no. 4 (1982), pp. 433-466.

18. Jobson, J.D., and B. Korkie, “On the Jensen Measure and Marginal Improvements in Portfolio Performance: A Note,” *Journal of Finance*, 39, no. 1 (1984), pp. 245-251.

19. Jobson, J.D., and B. Korkie, “A Performance Interpretation of Multivariate Tests of Asset Set Intersection, Spanning, and Mean-Variance Efficiency,” *Journal of Financial and Quantitative Analysis*, 24, no. 2 (1989), pp. 185-204.

20. Kan, R., and G. Zhou, “Mean Variance Spanning Tests,” working paper, University of Washington-St. Louis (2001).

21. Kodde, D.A., and F.C. Palm, “Wald Criteria for Jointly Testing Equality and Inequality Restrictions,” *Econometrica*, 54(5) (1986), pp, 1243-1248.

22. Landon, S., and C.E. Smith, "Government Debt Spillovers and Creditworthiness in a Federation," *Canadian Journal of Economics*, 33, no. 3 (2000), pp. 634-661.

23. Landon, S., and C.E. Smith, "Seasonality in Canadian Bond Returns: The Role of International Factors," *Canadian Journal of Administrative Sciences*, 23, no. 4 (2006), pp. 352-368.

24. Landon, S., and C.E. Smith, "Government Debt Spillovers in a Monetary Union," *North American Journal of Economics and Finance*, 18 (2007), pp. 135-154.

25. Markowitz, H., "Portfolio Selection," *Journal of Finance*, 7, no. 1, (1952), pp. 77-91.

26. Newey, W.K., and K.D., West, "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, no. 3(1987), pp. 703-708.

27. Stewart, K.G., "The Functional Equivalence of the W, LR, and LM Statistics," *Economics Letters*, 49, no. 2 (1995), pp. 109-112.

28. Van Zijl, T., "Risk Decomposition: Variance or Standard Deviation - A Reexamination and Extension," *Journal of Financial and Quantitative Analysis*, 22, no. 2 (1987), pp. 237-247.