# Is portfolio home bias pathological or an optimal response?

by

Adam Cooper

An essay submitted to the Department of Economics

in partial fulfillment of the requirements for

the degree of Master of Arts

Queen's University

Kingston, Ontario, Canada

July 2009

## Acknowledgements

I would like to thank Professor Gregor Smith for his patience, thoughtful suggestions, and time. Professor Smith's dedication to his students, his research interests, and his teaching fostered my interests in finance and econometrics and inspired much of the work herein. I would also like to thank Professor Morten Nielsen whose passion for econometrics, particularly time series, provided sound foundations and an impetus for my own empirical forays into the subject. I would also like to thank Professor Frank Milne whose course piqued my interest in financial economics. Also, David Byrne and Lee Morin — two of the best teaching assistants I have encountered — provided feedback, help, and support throughout the year well beyond what was required of them.

My instructors at McMaster University provided an excellent foundation for future work in economics. I am particularly indebted to Professor Atif Kubursi who helped spark my interest in economics and provide guidance throughout my undergraduate career. I am also thankful to have had Professors William Scarth, Frank Denton and Seungjin Han, all of whom fostered my interest in economics.

I also want to thank my colleagues who have been so generous in their willingness to help me work through many of the problems I experienced, both with respect to this essay and school work in general, particularly Patrick Day, Daniel Kou, Scott Robson, Terry Lio, and Curtis Chow.

I greatly appreciate the help and support I have received throughout the year from the Economics Department's staff, particularly Rachel Ferreira.

Lastly, I would like to thank the Economics Department, Queen's University, and the Government of Ontario for their provision of generous funding.

All errors and omissions are my own.

## Contents

## List of Tables



## List of Figures

# 1 Introduction

A very prominent puzzle in macroeconomics is the substantial preference international investors have for their respective home market. To most researchers, the tendency for investors to hold most of their assets in the domestic market is inefficient given the well-known gains from diversification. Does this tendency imply that investors are not acting rationally, or are there other explanations? For instance, could it be that studies which have found that home bias is inefficient are simply wrong? Specifically, is there a realistic model in which stock prices may be strongly related in the long-run and thus mitigate possible gains from diversification? If so, do empirical results indicate that stocks are highly related in the long-run and thus substantiate the model? Of course, if there is a model of highly related long-run stock prices and supporting evidence, the implication is that the cost of home bias is less than currently thought, and thus, that the home bias puzzle is not quite so puzzling.

To examine the possibility of the scenario above, I develop a bivariate random walk model of stock prices with a parameter which determines the weight on the shared stochastic trend (*i.e.,* at one extreme the two prices are cointegrated, while at the other they have no shared stochastic trends). With the implications of the model in mind, I then analyze how theoretical investors with varying investment horizons change their portfolio weights between domestic and international markets. I then use total return indices for the national markets of the G7 nations to test if there are any pairwise cointegrating vectors and if there are any cointegrating vectors in a systems framework.

My theoretical results indicate that as the weight on the shared stochastic trend increases, investors show a tendency to weight their portfolios toward the asset with the greater risk-reward tradeoff (*i.e.,* investors do not diversify their asset holdings as much when prices are cointegrated), which magnifies the effect of the differing perceptions hypothesis of French and Poterba (1991). While the theory indicates that in the presence of cointegrated markets, home bias is not as puzzling, my empirical results indicate that only a handful of the G7 nations share stochastic trends. In other words, there should still be substantial gains from diversification — even in the long-run — but less than perhaps was originally thought.

## 2 Research Context

In finance, the gains from portfolio diversification have long been known. Indeed, over forty years ago Grubel (1968) provided sound theoretical foundations for the examination of the gains (or lack thereof) from international diversification. He did so by expanding the widely accepted portfolio models of Markovitz (1959) and Tobin (1958) to incorporate both domestic and foreign asset holdings (*i.e.,* he extended the analysis from different assets to different countries). Empirically, using realized correlations between national stock indices, Grubel (1968) showed that holding foreign assets as opposed to a wholly domestic portfolio is beneficial in two unique ways: (i) income (and consumption) is decoupled from national business cycles, and (ii) productivity is increased due to improved resource flows. Levy and Sarnat (1970), Grubel and Fadner (1971), Ripley (1973), and Panton *et al.* (1976) all expanded on the work of Grubel (1968), and provided further evidence that gains from diversification existed internationally because markets were not perfectly correlated over time. Indeed, they found that this held both in the short-run and the long-run for the periods they examined.

While these early studies were innovative and important, they were carried out during a period marked by relatively high barriers to international capital flows. However, much has changed in the four decades since these original studies were completed. Indeed, in the forty years since then, trade barriers have consistently and substantially decreased, financial markets have become increasingly integrated, and information flows have become faster and cheaper. Thus, one would expect that over this long period of time, investors would have, at the very least, began to take advantage of the easier access to foreign markets and diversified their portfolios internationally. Strangely, many observers have discovered that this is not the case. Indeed, researchers have repeatedly found that investors do not engage in much international diversification at all.

In an important paper, French and Poterba (1991) examined both the extent to which investors exhibit a predilection to invest in their home markets and the cost to investors of doing so. Drawing on their 1990 paper (see French and Poterba, 1990) they noted that if the pairwise correlations between the returns from two different equity markets are less than one, then there exist gains from international diversification. Thus, they examined such pairwise correlations between the United States, Japan, the United Kingdom, France, Germany, and Canada and found that the correlation coefficients for each possible pair were far from one — the average correlation was 0.502 over the period of 1975-1989 — implying diversification would reduce risk. In other words, if investors own assets in foreign markets they can reduce their risk exposure without necessarily reducing their expected returns. Next, French and Poterba (1991) considered the optimal portfolio weights for investors with logarithmic constant absolute-risk-aversion (CARA) utility functions defined over wealth. Problematically, to solve for the optimal weights, one needs to estimate mean returns which cannot be estimated consistently. Thus, rather than estimate a parameter inconsistently, they utilized data on the portfolio weights investors had assigned and then "backed out" the implied average return investors must have been expecting over the period. The results of French and Poterba (1991) indicate a very strong preference for home markets, with the authors noting "that current portfolio patterns imply that investors in each nation expect returns in their domestic equity market to be several hundred basis points higher than returns in other markets." The looming question left unanswered by French and Poterba (1991) is "Why are investors not taking advantage of the international diversification opportunities presented by increasing financial market integration and information flows?" Alternatively, why does Lewis (1999) seem to believe that investors do not to a good job of "hedging risks across countries?"

These questions stem from a deeper and not very well understood issue known as home bias. Indeed, Obstfeld and Rogoff (2000) list home bias as the first of their six major puzzles in international macroeconomics. While it is true that Obstfeld and Rogoff (2000) were considering home bias with respect to international trade, the issues they raised are very similar to those of French and Poterba (1991): borders play rather shockingly large roles in the way individuals allocate resources. Moreover, the role of borders is even more puzzling in financial markets because they do not require the movement of physical goods or services like trade often does.

Thus, in order to appropriately examine and possibly unravel the puzzle of portfolio home bias, it seems pertinent to consider whether it and a lack of consumption risk sharing (the topic Obstfeld and Rogoff (2000) puzzled over) can be thought of individually, or are actually two symptoms of deeper issues. Recently, some studies analyzed the relationship between the two and found that perhaps they are strongly related. For example, Sørensen *et al.* (2007) provided evidence that consumption risk sharing and investor home bias are actually "closely related empirical phenomena" when examined within the Organization for Economic Development and Cooperation (OECD hereafter). Also, the authors noted that while portfolio home bias still exists, it appeared to have declined between 1993 and 2003, the period of their study. Coeurdacier (2009) expanded on the findings of Sørensen *et al.* (2007) and developed a stochastic equilibrium model in which trade costs lead to a home bias in portfolios. Indeed, Coeurdacier (2009) provided a very elegant but reasonable interpretation of his results: as the cost of trading goods increases, domestic income volatility falls due to "softened competition." The result is that there is less risk in the domestic equity market and thus investors have fewer incentives to diversify internationally. In short, portfolio home bias and a lack of international risk sharing are closely related. Indeed, a lack of international risk sharing may actually lead to a rational home bias in portfolios. While some studies contest the rationality of home bias (*e.g.,* Karlsson and Nordén (2007) find that the opposite may be the case, that the "likelihood of home bias is caused by both rational and irrational factors") it is clear that there is a large body of evidence indicating that portfolio home bias and a lack of consumption risk sharing are strongly related.

Adding to the research on the relationship between the two phenomena, Lewis (1996) found that the lack of international consumption risk sharing was likely partially explained by capital market controls (much like the findings by Sørensen *et al.* (2007) and Coeurdacier (2009) who found that consumption risk sharing and investor home bias are closely related). Could capital controls also explain the lack of international risk sharing present in worldwide equity markets? Kho *et al.* (2006) found that perhaps this is the case. In particular, while they found that if countries are equally weighted, home bias did not fall between 1994 and 2004 — a period in which many barriers to international trade declined — if countries are weighted instead based on their market capitalization, home bias did actually decrease between 1994 and 2004. To explain this finding, Kho *et al.* (2006) concluded that it may not actually be a preference for home markets that leads to home bias, but rather excessive "insider" ownership of firms. Specifically, since foreigners cannot typically own shares which are open only to insiders, a market which has a high fraction of domestic insiders is likely to have low levels of foreign investment. In summary, many studies indicate that portfolio home bias and a lack of consumption risk sharing are closely related phenomena which researchers examining either should bear in mind.

Interestingly, as noted above, many studies have found that home bias is declining. For example, Baele *et al.* (2007) found that under alternative frameworks (*i.e.,* distinct from the I-CAPM) home bias has decreased for many countries since the early 1990s, but by no means disappeared. Moreover, Baele *et al.* (2007) also found persuasive evidence "that globalization and regional integration, and especially its most intense form in the Euro Area, relate significantly to the decrease of home bias."

While home bias in terms of both consumption risk sharing and investors' portfolio holdings is declining, it is nonetheless still a major puzzle in macroeconomics, with many studies still examining different aspects of the dominance of domestic markets in investor portfolios. Home bias also is the primary impetus for this study; however, the main thrust of my research is toward determining whether or not a home bias in portfolios is *pathological* or actually an *optimal* response to less than otherwise expected long-term gains from diversification.

As noted above, many of the early studies found that, at least in the 1960s, there did exist both short and long term gains from international diversification. However, there are many pitfalls associated with the methods used by the authors. Indeed, Allen and MacDonald (1995) provide a succinct summary of the problems associated with the earlier studies by Grubel *et al.*: "[they] used *ex post* analysis in which it is assumed that the required inputs (expected returns, variances and covariances) estimated to form internationally diversified portfolios are known with *certainty*." The problems of estimating covariance structures were exacerbated further by the often very short time horizon of their studies due to data limitations (sometimes less than five years). Furthermore, as previously mentioned, since most early studies focused on *ex post* correlations (or covariance structures in the case of

Ripley (1973)) the econometric method may have actually been wrong. Indeed, since the theoretical development of cointegration by Granger (1981), it has become clear that using econometric techniques intended for stationary variables with non-stationary variables involves a mispecification. As Kasa (1992) notes, most early studies of international diversification opportunities used return data rather than levels (of prices or indices), which, according to Engle and Granger (1987), is at best inefficient, and at worst, biased. Thus, almost all recent studies which focus on either the possible gains from international diversification or national capital market integration utilize cointegration techniques with respect to the logged levels of indices (for examples, see Chang (2001), Taylor and Tonks (1989), Narayan and Smyth (2004), Kanas (1998), and Kanas (1999)).

The use of cointegration techniques allows for much greater flexibility in examining whether or not there exist gains from international diversification. In particular, while using correlations or covariances allows the researcher to (at best) examine the possibility that there are any gains from diversification, cointegration allows the researcher to not only test if gains from diversification exist, but also determine if they differ between the short- and long-run and provide insights into what may be considered "long-term." For example, while stock prices may appear unrelated over a short time horizon, there is a distinct possibility that they are highly related in the long-run due to a shared stochastic trend. Indeed, this allows cointegration analysis to indicate whether diversification opportunities exist for time horizons that are *longer* than the study period, whereas correlations and covariances, because such calculations do not consider the existence of stochastic trends, are only valid for the time period considered.

Now, consider the case where two markets share a stochastic trend. Under this condition, it *may* be understandable for long term investors to maintain much of their portfolio in domestic markets, rather than split their holdings between their home markets and foreign ones because all markets may be highly related and there may be added risk and/or taxes when investing abroad. In other words the home bias of investors, in this environment of shared stochastic trends, is an optimal response, and thus the macroeconomic puzzle of "home bias" is less of a conundrum. To explore this possibility, I extend the theoretical analysis and also update the empirical evidence to examine the potential for G7 investors to gain from international diversification within the G7 itself.

In the theoretical sections I explain and justify my selection of a *random walk* process for asset prices. Then, under this specification, I develop a unique bivariate model to investigate four issues: (i) exactly how cointegration in *random walk* asset prices may affect unconditional covariance structures, (ii) how investors would change their portfolios in the presence or absence of cointegrated asset prices (I also examine the intermediate case where there is a common stochastic trend but the asset prices are *not* cointegrated), (iii) how cointegrated prices would affect home biases in an environment with varying perceptions of mean returns and risk, and (iv) how the investor's time horizon affects his portfolio allocation problem.

Importantly, my research indicates that cointegration does indeed matter to investors, and in the presence of a shared stochastic trend, the gains from diversification are significantly lower than when the asset prices have no common stochastic trends. Moreover, under my specification, the more closely related the two prices in terms of their permanent shocks, the more polarized the asset holdings will be (the balance of holdings leans primarily toward on the asset with the higher return). Moreover, by utilizing numerical examples and the differing perceptions hypothesis of French and Poterba (1991), I illustrate how cointegrated markets would magnify the impact of different expectations on home bias (*i.e.,* investor expectations of domestic returns may not be as high as French and Poterba (1991) found). While the polarization of portfolios is an increasing function of the time horizon, the result nonetheless holds for even single period time horizons. Furthermore, the results hold under both the traditional portfolio problem and a CARA utility function as used by French and Poterba (1991). Not surprisingly, I also find that the degree to which the asset prices share permanent shocks becomes increasingly important as the investor's time horizon increases, similar to the findings of other researchers.

In the empirical portion (Sections 6 to 9) of this essay, I examine the possibility of shared stochastic trends using the *total* return indices constructed and maintained by MSCI Barra for Canada, France, Germany, Italy, Japan, the United Kingdom and the United States (*i.e.,* the G7) over the period from January 1999 to July of 2009. These indices have been converted to US dollars to provide a common base for comparison. The frequency of observation is weekly. Since MSCI Barra produces both net and gross total return indices for all G7 nations, both are utilized for robustness (see Section 6 for definitions). After selecting the data, I first test the random walk hypothesis using both augmented Dickey and Fuller (1979) and Phillips and Perron (1988) tests and find that I cannot reject the null hypothesis of a unit root for any of the series. Moreover, I also find that after first differencing the data, each series is stationary implying that all are integrated of order one and thus that cointegration analysis is appropriate. I then test for cointegration using both a bivariate and a systems approach. While a bivariate approach my seem simplistic at first because investors are, of course, able to invest in multiple countries at a time, a pairwise specification is useful because there are still important economic implications, it allows for multiple robustness checks, and avoids some of the problem associated with a systems analysis.

To test for pairwise cointegration I use three different testing techniques: the Engle and Granger (1987) two-step, the Zivot and Andrews (1992) test allowing for a structural break, and the Johansen (1988) procedure. With the Zivot and Andrews (1992) tests I find five out of 21 possible cointegrating relationships. The other two procedures indicate even fewer cointegrating relationships. In other words, it does not appear to be the case that over the decade beginning in 1999 that the equity markets of the G7 have become highly related. Instead, it appears that long-run gains from diversification are still present. Strikingly, the United States is involved in almost all of the pairwise cointegrating relationships which suggests that it would be best for long-term investors to focus on non-US markets to maximize their gains from diversification. In terms of the system as a whole, the results are similar for both indices. Specifically, I find no cointegrating vectors in terms of the gross indices or net indices after the Reimers (1992) finite sample correction is applied. In other words, when analyzing the system as a whole, there appear to be large gains from diversification within the G7.

## 3 Random Walk Prices and Cointegration

### *3.1 Why a random walk?*

The most crucial assumption of this paper is that stock prices are appropriately modeled by random walks. Such a specification of course indicates that stock prices are simply accumulations of random shocks, and thus that previous changes do not help predict future changes. In that sense, one can think of each change in the stock price (aside from a trend term) as being *uncorrelated* with previous changes. Indeed, in an extremely influential (and cited) study, Fama (1965) found that for the thirty stocks in the Dow Jones Industrial Average over the period of 1957 to 1962 there was "strong and voluminous evidence in favour of the random walk hypothesis." The random walk specification is somewhat contested, however. For example, studies like those of Lo and MacKinlay (1988) and Poterba and Summers (1989) concluded that stock prices are not best modeled as random walks. Moreover, there have been challenges to the application of unit roots to economic problem in general. Sims (1988), a particularly vociferous opponent of unit roots in economics, asserted that the distinction between stationary and unit root data is not fundamentally important. Cochrane (1991) and other researchers have similarly criticized studies focused on unit roots due to the well known size and power issues surrounding many of the related testing procedures. Lastly, as evidenced by the contradictory findings by Perron (1989) and Zivot and Andrews (1992) unit root tests are extremely sensitive to the decisions by the researcher as to whether or not structural breaks are endogenous, exogenous, or even present.

Given the above problems associated with unit roots and the unit root hypothesis of stock prices, I adopt the stance of Kasa (1992) *et al.*: while I do not contribute to the debate in this essay with respect to either unit root tests or the unit root hypothesis, I utilize the standard random walk tests developed by Dickey and Fuller (1979) and Phillips and Perron (1988) and, given that the results do indeed indicate the series are unit roots, utilize a random walk model of stock prices.

### *3.2 Random walk representation of prices*

Consider two assets with prices $A_t$ and $B_t$, and which are from two different countries (say, countries $A$ and $B$). The natural logarithm of their prices (hereafter log prices) are given by $a_t$ and $b_t$. Suppose now that $a_t$ and $b_t$ can be modeled as suggested in Section 3.1 as random walks given by:

$$a_t = \sum_{i=0}^{t} \varepsilon_{ai} + \eta_{at} + \mu_a t \tag{1}$$

$$b_t = \sum_{i=0}^{t} \varepsilon'_{bi} + \eta_{bt} + \mu_b t. \tag{2}$$

I assume that $\varepsilon'_{bt}$ can be decomposed into two components: $\varepsilon_{at}$ and $\varepsilon_{bt}$ where $\varepsilon_{bt}$ is orthogonal to $\varepsilon_{at}$. Specifically, $\varepsilon'_{bt}$ is given by:

$$\varepsilon'_{bt} = \beta\varepsilon_{at} + (1-\beta)\varepsilon_{bt}. \tag{3}$$

Thus, combining the log price (2) and the permanent shock decomposition (4), $b_t$ can be rewritten as:

$$b_t = \beta \sum_{i=0}^{t} \varepsilon_{ai} + (1-\beta) \sum_{i=0}^{t} \varepsilon_{bi} + \eta_b + \mu_b t. \tag{4}$$

In order to maintain as much generality and realism as possible, the covariance structure of $a_t$ and $b_t$ is somewhat complicated. In particular, while I assume that all shocks are uncorrelated between periods, I allow for correlation between the contemporaneous idiosyncratic shocks, $\eta_{at}$ and $\eta_{bt}$. Formally, I assume that $\varepsilon_{xt} \sim iiN(0, \sigma_{x\varepsilon})\ \ \forall\ x = a, b$; $\eta_{xt} \sim iiN(0, \sigma_x^2)\ \ \forall\ x = a, b$; $E(\eta_{at}\eta_{bt}) = \sigma_{ab}$; $E(\varepsilon_{at}\varepsilon_{bt}) = 0$; $E(\varepsilon_{xt}\eta_{zt}) = 0\ \ x = a, b$ and $z = a, b$; but that $E(\varepsilon_{at}\varepsilon'_{bt}) \neq 0$. The allowance for a covariance structure between idiosyncratic shocks (*i.e.,* non-permanent shocks) is made because there are often relationships between contemporaneous movements in stock prices. Also, I assume $\beta \in [0, 1]$. While there is not a technical reason that $\beta$ must be positive, a negative value of $\beta$ would imply negative correlation between $a_t$ and $b_t$ which is economically uninteresting. Thus, assuming that $\beta$ is always positive is a simplifying and realistic condition. Lastly, in order to make the problem tractable from an investor's perspective, I have to assume a specific relationship between the variances of $\varepsilon_{at}$ and $\varepsilon_{bt}$. I outline the argument and derivation below in Section 3.4.

Equation (4) provides an intuitive illustration of the relationship between $a_t$ and $b_t$: the greater the value of $\beta$ the more closely related the two prices. Indeed, if $\beta = 1$, $a_t$ and $b_t$ are cointegrated. This result is explained below in Section 3.3.

Importantly, both prices are *non-stationary* (*i.e.,* the variances of $a_t$ and $b_t$ are time-dependent). However, consider the first differences of both:

$$\Delta a_t = \varepsilon_{at} + \Delta\eta_{at} + \mu_a \tag{5}$$

$$\Delta b_t = \beta\varepsilon_{at} + (1-\beta)\varepsilon_{bt} + \Delta\eta_{bt} + \mu_b \tag{6}$$

where $E(\Delta a_t) = \mu_a$ and $E(\Delta b_t) = \mu_b$. Clearly, the first differences are stationary. This indicates that the log prices are integrated of order one, hereafter $I(1)$ (the order of integration implies the number of times the data must be differenced in order to be left with a stationary variable).

### *3.3 Cointegration and prices*

While variables that are *non-stationary* are problematic in a traditional regression framework (*i.e.,* such variables lead to spurious regressions in general), they are somewhat easily dealt with in a cointegration environment. Moreover, the parameters estimated using cointegration techniques have very useful interpretations. For example, using the Engle and Granger (1987) two-step procedure, the estimated slope parameter indicates the *sensitivity* of the relationship.

In the environment above, defined by equations (1) and (4), suppose $\beta = 1$ and consider the linear combination:

$$[1 \quad -1][b_t \quad a_t]' = \eta_{bt} - \eta_{at}. \tag{7}$$

Such a linear combination gives a stationary $I(0)$ variable. This relationship is what defines cointegration: there exists a linear combination of non-stationary variables that is itself stationary. In the example above, $a_t$ and $b_t$ are cointegrated if and only if $\beta = 1$ (for any linear combination when $\beta \neq 0$ the result will not be stationary). While the restriction of

$\beta = 1$ is extreme, it will lead to relatively specific correlation structures between $a_t$ and $b_t$ which are investigated and described below.

### *3.4 Unconditional correlation*

Before I consider the consequences of the cointegration of asset prices on portfolios, I examine the unconditional correlation between prices $a_t$ and $b_t$. Intuitively, if the limit, as time tends to infinity, of the unconditional correlation between $a_t$ and $b_t$ is one, it indicates that (i) over the long-run, the stochastic trends drive the relationship, and (ii) there is likely to be little to gain from international diversification if the investor's time horizon is long. Thus, the unconditional correlation offers a convenient starting point for this essay. In order to evaluate it though, I need to calculate the unconditional variances of both prices and the unconditional covariance between them. Consider first the unconditional variance of $a_t$, given by:

$$V(a_t) = E\left[\left(\sum_{i=0}^{t} \varepsilon_{ai} + \eta_{at} + \mu_a t\right) - \mu_a t\right]^2 \tag{8}$$

$$= E\left(\sum_{i=0}^{t} \varepsilon_{ai}^2\right) + E\left(\eta_{at}^2\right) = t\sigma_{a\varepsilon}^2 + \sigma_a^2. \tag{9}$$

Similarly, the unconditional variance of $b_t$ is given by:

$$V(b_t) = E\left[\left(\beta \sum_{i=0}^{t} \varepsilon_{ai} + (1-\beta) \sum_{i=0}^{t} \varepsilon_{bi} + \eta_{bt} + \mu_b t\right) - \mu_b t\right]^2 \tag{10}$$

$$= \beta^2 E\left(\sum_{i=0}^{t} \varepsilon_{ai}^2\right) + (1-\beta)^2 E\left(\sum_{i=0}^{t} \varepsilon_{bi}^2\right) + E\left(\eta_{bt}^2\right) \tag{11}$$

$$= t\left[\beta^2 \sigma_{a\varepsilon}^2 + (1-\beta)^2 \sigma_{b\varepsilon}^2\right] + \sigma_b^2. \tag{12}$$

As noted above, all shocks are uncorrelated between periods, and there is covariance between the idiosyncratic shocks only. Importantly, the variances depend on time. Intuitively, since $a_t$ and $b_t$ are both the sums of shocks, the variance in either depends on how many shocks occur over the period of observation.

From an intuitive standpoint, one aspect of equation (12) is particularly troubling, however. Specifically, $\beta$ has two important roles in the definition of $b_t$: (i) it indicates the strength of the relationship between $a_t$ and $b_t$, and (ii) it strongly affects the variance of $\varepsilon'_{bt}$. It is the second aspect that is problematic — indeed, the first aspect is exactly what $\beta$ is intended for — and must be addressed in order to make the problem tractable.

To ameliorate the affect $\beta$ has on the variance (12), I begin from the standpoint that $\beta$ should *not* affect the variance of $\varepsilon'_{bt}$. To do so, I set the variance of the first stochastic trend equal to that of the second. Using this assumption, along with $V(\varepsilon'_{bt}) = \beta^2\sigma^2_{a\varepsilon} + (1-\beta)^2\sigma^2_{b\varepsilon}$, gives:

$$\sigma^2_{a\varepsilon} = \beta^2\sigma^2_{a\varepsilon} + (1-\beta)^2\sigma^2_{b\varepsilon} \tag{13}$$

$$(1-\beta^2)\sigma^2_{a\varepsilon} = (1=\beta)^2\sigma^2_{b\varepsilon} \tag{14}$$

$$\sigma^2_{b\varepsilon} = \frac{(1-\beta^2)\sigma^2_{a\varepsilon}}{(1-\beta)^2}. \tag{15}$$

To summarize, the variance of the second stochastic trend varies so that a change in $\beta$ does not affect the variance of $b_t$. Interestingly, equation (15) implies that $\sigma_{b\varepsilon}$ approaches infinity as $\beta$ approaches one. This is a necessary result, however, in order for $V(\varepsilon'_{bt}) = V(\varepsilon_{at})$ to hold.

Now, combing equation (15) with the variance of $b_t$ (12) gives:

$$V(b_t) = t\left[\beta^2\sigma^2_{a\varepsilon} + (1-\beta)^2\left(\frac{(1-\beta^2)\sigma_{a\varepsilon}}{(1-\beta)^2}\right)\right] + \sigma^2_b \tag{16}$$

$$= t\sigma^2_{a\varepsilon} + \sigma^2_b. \tag{17}$$

Comparing $V(a_t)$ to $V(b_t)$, it is clear that the stochastic trends now have the same variance, and, more importantly, that varying $\beta$ has *no* effect on the variance of $b_t$. Moreover, I can now simplify the notation since $V(\varepsilon_{bt}) = V(\varepsilon_{at})$, *i.e.,* I now define $\sigma^2_\varepsilon \equiv \sigma^2_{a\varepsilon} = \sigma^2_{b\varepsilon}$ and use this notation for the rest of the paper. Accordingly, I can now rewrite variances (9) and (12) using the new notation, as:

$$V(a_t) = t\sigma^2_\varepsilon + \sigma^2_a \tag{18}$$

$$V(b_t) = t\sigma^2_\varepsilon + \sigma^2_b. \tag{19}$$

The last formula required for the unconditional correlation between $a_t$ and $b_t$ is the unconditional covariance, which is defined as:

$$Cov(a_t, b_t) = E(a_t - E(a_t))(b_t - E(b_t)). \tag{20}$$

Thus, the unconditional covariance is given by:

$$Cov(a_t, b_t) = E\left(\sum_{i=0}^{t} \varepsilon_{ai} + \eta_{bi}\right)\left(\beta \sum_{i=0}^{t} \varepsilon_{ai} + (1-\beta)\sum_{i=0}^{t} \varepsilon_{bi} + \eta_{bt}\right) \tag{21}$$

$$= t\beta\sigma_\varepsilon^2 + \sigma_{ab}. \tag{22}$$

Note that the correlation (22) depends on time. Again, the existence of a time element reflects of the definition of $a_t$ and $b_t$ as the sums of random shocks. Lastly, note that while it was problematic that $\beta$ affected the variance of $b_t$, it is appropriate that it affect the covariance between $a_t$ and $b_t$, since $\beta$ is intended to be a measure of the relationship between the two processes. Indeed, as time increases, so long as $\beta > 0$, the stochastic trend, $\varepsilon_{at}$, drives the relationship.

Now, having calculated the variances of both prices and their covariance terms I can derive the unconditional correlation between $a_t$ and $b_t$:

$$Corr(a_t, b_t) = \frac{cov(a_t, b_t)}{\sqrt{V(a_t)}\sqrt{V(b_t)}} = \frac{t\beta\sigma_\varepsilon^2 + \sigma_{ab}}{\sqrt{t\sigma_\varepsilon^2 + \sigma_a^2}\sqrt{t\sigma_\varepsilon^2 + \sigma_b^2}}. \tag{23}$$

To examine the asymptotic properties of equation (23) I first divide and multiple both the numerator and denominator by a factor of $t$, giving:

$$Corr(a_t, b_t) = \frac{t(\beta\sigma_\varepsilon^2 + \frac{\sigma_{ab}}{t})}{t\left(\sqrt{\sigma_\varepsilon^2 + \frac{\sigma_a^2}{t}}\sqrt{\sigma_\varepsilon^2 + \frac{\sigma_b^2}{t}}\right)}. \tag{24}$$

Clearly, the $t$ factors outside the parenthesis cancel, leaving only three terms containing $t$. Since the limit of a product is equal to the product of the limits I evaluate the limits of each term individually. Consider first the limit of the numerator, which is given by:

$$\lim_{t\to\infty} \beta\sigma_\varepsilon^2 + \frac{\sigma_{ab}}{t} = \beta\sigma_\varepsilon^2 \tag{25}$$

which is independent of time. Similarly, consider the limits of the two terms in the denominator, which are given by:

$$\lim_{t\to\infty} \sigma_\varepsilon^2 + \frac{\sigma_a^2}{t} = \sigma_\varepsilon^2 \tag{26}$$

$$\lim_{t\to\infty} \sigma_\varepsilon^2 + \frac{\sigma_b^2}{t} = \sigma_\varepsilon^2. \tag{27}$$

Again, the limits in equations (26) and (27) are independent of time. Now, using equations (25) to (27), the limit of equation (24) reduces to:

$$Corr(a_t, b_t) = \frac{\beta\sigma_\varepsilon^2}{\sqrt{\sigma_\varepsilon^2}\sqrt{\sigma_\varepsilon^2}}. \tag{28}$$

Importantly, if $\beta = 1$ then $a_t$ and $b_t$ are cointegrated and equation (28) simplifies to:

$$Corr(a_t, b_t) = \frac{\sigma_\varepsilon^2}{\sqrt{\sigma_\varepsilon^2}\sqrt{\sigma_\varepsilon^2}} = 1. \tag{29}$$

That is, if $a_t$ and $b_t$ are cointegrated, as $t \to \infty$, $Corr(a_t, b_t) \to 1$. In other words, the shared stochastic trend dominates the relationship. Conversely, if $\beta = 0$ then the unconditional correlation coefficient is zero even without using asymptotics. If $\beta \in (0, 1)$ the correlation coefficient is similarly between 0 and 1. Note also that the lower bound of the unconditional correlation (29) is 0 since $\beta$ is assumed to be positive (a negative correlation is unrealistic). The result in equation (29) is very important in this essay because it indicates that if the prices of two assets are cointegrated there are quite possibly few long-run gains from diversification. DeFusco *et al.* (1996) similarly note that in the case of cointegrated variables (recall that this is equivalent to $\beta = 1$), the idiosyncratic shocks "die out" and thus allow the shared stochastic trend to drive the relationship.

## 4 Returns

### *4.1 Unconditional one-period returns*

The main focus of investors is, of course, not the price level of an asset, but the return earned. Recall that $a_t$ and $b_t$ are defined as the logs of the prices $A_t$ and $B_t$ respectively. Thus, the returns for each asset can be approximated well by the first difference of both log prices, *i.e.*, $r_{at} \simeq \Delta a_t$ and $r_{bt} \simeq \Delta b_t$.

Notice, however, that both first differences, $\Delta a_t$ (5) and $\Delta b_t$ (6), contain first-order moving averages with respect to the temporary shocks, implying that the returns are partly predictable. The predictability in turn indicates that the conditional returns will be slightly different than the unconditional ones. For example, let $\Omega_t$ be the information set at time $t$. The expected return of asset $A_t$ conditional on the information set at time $t-1$ is given by:

$$E(r_{at}|\Omega_{t-1}) = E(\varepsilon_{at} + \eta_{at} - \eta_{a,t-1} + \mu_a|\Omega_{t-1}) \neq E(r_{at}) = \mu_a. \tag{30}$$

The conditional return (30) is clearly negatively correlated with the previous period's temporary shock, $\eta_{a,t-1}$.

While in theory describing behaviour based on the conditional returns may be more realistic, I focus on the more tractable unconditional returns for two reasons. First, the predictability in the returns would be hard to detect in the noise of stock prices (recall that random walks have unpredictable changes aside from the drift term). Second, in the application $t$ counts weeks, so any predictability will vanish after one week. It seems unreasonable to assume that investors (particularly long-term investors) will rebalance their portfolios weekly. Moreover, since I use the unconditional returns which are constant over time, so too are the optimal portfolio weights.

Accordingly, I now examine the unconditional variance-covariance structure of one-period returns for $a_t$ and $b_t$ and use the results to investigate the effects of increasingly related log prices on a portfolio containing both assets. Thus, a portfolio, $C_t$, with a fraction $g$ invested in asset $A_t$ and a fraction $(1-g)$ invested in asset $B_t$ (*i.e.*, $g \in [0,1]$)

will have an unconditional expected return given by:

$$E(r_{ct}) = gE(r_{at}) + (1-g)E(r_{bt}) = gE(\Delta a_t) + (1-g)E(\Delta b_t) \tag{31}$$

$$= g\mu_a + (1-g)\mu_b. \tag{32}$$

The variance of the portfolio is simply the variance of a linear combination of random returns. Thus, the variance of the portfolio is given by:

$$V(r_{ct}) = g^2V(r_{at}) + (1-g)^2V(r_{bt}) + 2g(1-g)Cov(r_{at}, r_{bt}). \tag{33}$$

Accordingly, to determine the variance of the portfolio, I need to calculate the variances of $r_{a,t}$ and $r_{b,t}$, and the covariance between them. First, consider the variance of the returns, given by:

$$V(r_{a,t}) = E(\varepsilon_{at} + \Delta\eta_{at} + \mu_a - \mu_a)^2 = \sigma_\varepsilon^2 + 2\sigma_a^2 \tag{34}$$

$$V(r_{b,t}) = E(\beta\varepsilon_{at} + (1-\beta)\varepsilon_{bt} + \Delta\eta_{bt} + \mu_b - \mu_b)^2 = \sigma_\varepsilon^2 + 2\sigma_b^2. \tag{35}$$

Now consider the theoretical covariance between the two returns, which is given by:

$$Cov(r_{a,t}, r_{b,t}) = E(\varepsilon_{at}+\Delta\eta_{at}+\mu_a-\mu_a)(\beta\varepsilon_{at}+(1-\beta)\varepsilon_{bt}+\Delta\eta_{bt}+\mu_b-\mu_b) = \beta\sigma_\varepsilon^2+2\sigma_{ab}. \tag{36}$$

Lastly, consider the variance of the portfolio given in equation (33) using equations (34) to (36):

$$V(r_{ct}) = g^2(\sigma_\varepsilon^2 + 2\sigma_b^2) + (1-g)^2(\sigma_\varepsilon^2 + 2\sigma_b^2) + 2g(1-g)(\beta\sigma_\varepsilon^2 + 2\sigma_{ab}). \tag{37}$$

The crucial aspect of the return covariance (36) is that while the return is a simple weighted average of the two individual returns, the variance is, in general, less than just a weighted average of the two individual variances. Indeed, it is well known in financial economics that gains from diversification exist whenever two or more assets are not perfectly positively correlated. The variance of the portfolio (37) provides a very concise illustration of such gains from diversification, since it is clear that as $\beta$ increases, the relationship between $a_t$ and $b_t$ becomes stronger and the variance of the portfolio also increases.

Herein lies the issue at the heart of this paper: are returns actually more correlated over time than otherwise thought? Similarly, might it be the case that as financial markets have become increasingly integrated the opportunities for long-term international investors to gain from diversifying their asset holdings are less than they may appear. The variance of the portfolio (37) indicates that this may be the case and, therefore, provides the foundation for this paper. Next, I consider the variance of the portfolio as the investor's time horizon increases.

### *4.2 Unconditional returns over longer periods*

The analysis done in Section 4.1 applies only when the return is defined as the first difference of the asset's log price (*i.e.,* $a_t$ or $b_t$); however, as determined in Section 3.4, the correlation of cointegrated variables tends to one only as time tends to infinity. That is to say that while in the short run the idiosyncratic shocks $\eta_{at}$ and $\eta_{bt}$ allow for gains from diversification (*i.e.,* it should not be surprising that investors will hold both assets even if they are cointegrated if they have a short time horizon), in the long-run the stochastic trends drive the system and eliminate most, if not all, gains from diversification. Thus, I now consider returns between periods $t$ and $t+n$. To do so, I first define the $n$-period returns on $a_t$ and $b_t$ as $r_{a,t+n} \equiv (1-L^n)a_{t+n}$ and $r_{b,t+n} \equiv (1-L^n)b_{t+n}$, respectively. Using the definitions of the log prices (1) and (4), the $n$-period returns are given by:

$$E(r_{a,t+n}) = E(a_{t+n} - a_t) = E\left(\sum_{t+1}^{t+n} \varepsilon_{ai} + (t+n)\mu_a + \eta_{a,t+n} - \eta_{at}\right) = n\mu_a \tag{38}$$

$$E(r_{b,t+n}) = E(b_{t+n} - b_t) = E\left(\beta\sum_{t+1}^{t+n} \varepsilon_{ai} + (1-\beta)\sum_{t+1}^{t+n} \varepsilon_{bi} + (t+n)\mu_b + \eta_{b,t+n} - \eta_{b,t}\right) = n\mu_b \tag{39}$$

While the $n$-period returns (38) and (39) are informative, the use of an $n$ period lag operator, $(1-L^n)$, rather than the first difference operator, does lead to a decrease in the accuracy of the log approximation of the return. However, it simultaneously and more importantly introduces a time variable which can provide intuition as to how returns are correlated between cointegrated variables over increasingly long investment horizons. Also, note that if $n = 1$, the $n$-period returns (38) and (39) are the same as those found in the one period returns (5) and (6). As in Section 4.1, I first find the variance-covariance structure of the returns and then examine the correlation coefficient for $r_{a,t+n}$ and $r_{b,t+n}$.

The variances of $r_{a,t+n}$ and $r_{b,t+n}$ are given by:

$$V(r_{a,t+n}) = E\left(\sum_{t+1}^{t+n} \varepsilon_{ai} + \eta_{b,t+n} - \eta_{bt}\right)^2 = n\sigma_\varepsilon^2 + 2\sigma_a^2 \tag{40}$$

$$V(r_{b,t+n}) = E\left(\beta\sum_{t+1}^{t+n} \varepsilon_{ai} + (1-\beta)\sum_{t+1}^{t+n} \varepsilon_{bi} + \eta_{b,t+n} - \eta_{bt}\right)^2 = n\sigma_\varepsilon^2 + 2\sigma_b^2. \tag{41}$$

In this case, since the return is defined over $n$ periods, the variances increase as $n$ increases. Importantly though, the additional variance added each extra period the return is defined over accrues only due to the permanent shocks (*i.e.,* the stochastic trends).

The covariance of the one period returns (36) is easily extended to the $n$-period return case, giving:

$$Cov(r_{a,t+n}, r_{b,t+n}) = E\left(\beta\sum_{t+1}^{t+n} \varepsilon_{ai}^2 + \eta_{a,t+n}\eta_{b,t+n} + \eta_{at}\eta_{bt}\right) = n\beta\sigma_\varepsilon^2 + 2\sigma_{ab}. \tag{42}$$

Now, using the correlation definition (23) with equations (40) to (42), the correlation coefficient of the multiple period returns is given by:

$$Corr(r_{a,t+n}, r_{b,t+n}) = \frac{n(\beta\sigma_\varepsilon^2 + \frac{2\sigma_{ab}}{n})}{n\left(\sqrt{\sigma_\varepsilon^2 + \frac{2\sigma_a^2}{n}}\right)\left(\sqrt{\sigma_\varepsilon^2 + \frac{2\sigma_b^2}{n}}\right)} \tag{43}$$

Again, if $n = 1$, the correlation for the $n$-period returns (43) is the same as the correlation for the one period returns, and is nearly identical to the correlation of $a_t$ and $b_t$ — although in the case of the asset prices the correlation was from $t = 0$ to $t = T$, whereas in the case of the returns, the period is from $t = T$ to $t = T + n$. Importantly, the correlation of the returns now explicitly incorporates the investment horizon. The fundamental point is that if the two prices are cointegrated, the correlation coefficient tends to unity as the investment horizon tends towards infinity.

In summary, without addressing portfolio theory directly (I do this in Section 5), I have now shown that if the log prices of two assets are cointegrated then: (i) the unconditional correlation coefficient tends to one, (ii) the covariance of the one-period returns is positively related to $\beta$, and (iii) as the investment horizon increases (*i.e.,* as $n$ increases), so too does the correlation of the $n$-period returns.

## 5 Investor Objective Functions and Optimal Allocations

### *5.1 The traditional portfolio problem*

The traditional portfolio problem for risk averse investors nonetheless provides an excellent starting point to examine how investors may behave when prices are or are not cointegrated. Recall that I use the unconditional returns derived in Sections 4.1 and 4.2. Now, consider the general problem of an investor wishing to maximize his expected return subject to his risk tolerance and the variance of the portfolio. The problem can be described as follows:

$$\max_{g} \quad E(r_{ct}) - \lambda V(r_{ct}). \tag{44}$$

The objective function (44) essentially indicates that an investor is willing to bear additional variance (*i.e.,* risk), so long as he is commensurately compensated with larger expected returns. The additional return required is indicated by $\lambda$ which is a measure of risk tolerance. I also do not include a time subscript on $g$ (portfolio weights) since at the optimum it will not be time dependent (it is, however, dependent on the investor's time horizon). The objective function (44) can be re-written as:

$$\max_{g} \quad gE(r_{a,t})+(1-g)E(r_{b,t})-\lambda\Big[g^2V(r_{a,t})+(1-g)^2V(r_{b,t})+2g(1-g)Cov(r_{a,t},r_{b,t})\Big]. \tag{45}$$

Finally, substituting in the expected return (32) and variance of the portfolio (37), the objective function (45) becomes:

$$\max_{g} \quad g\mu_a + (1-g)\mu_b - \lambda\Big[g^2(\sigma_\varepsilon^2 + 2\sigma_a^2) + (1-g)^2[\sigma_\varepsilon^2 + \sigma_b^2] + 2g(1-g)(\beta\sigma_\varepsilon^2 + 2\sigma_{ab})\Big]. \tag{46}$$

To maximize the objective function (46) I find the optimal portfolio weights, $g^*$.

To do so, I differentiate the objective function (46) with respect to $g$ and then set the partial derivative equal to zero to determine the optimal portfolio holdings. The first order condition is:

$$0 = \mu_a - \mu_b - \lambda\left[2g(\sigma_\varepsilon^2 + 2\sigma_a^2) - 2(1-g)[\sigma_\varepsilon^2 + 2\sigma_b^2] + (2-4g)(\beta\sigma_\varepsilon^2 + 2\sigma_{ab})\right]. \tag{47}$$

After rearranging, the optimal value, $g^*$, is given by:

$$g^* = \frac{\mu_a - \mu_b + 2\lambda[\sigma_\varepsilon^2(1-\beta) + 2\sigma_b^2 - 2\sigma_{ab}]}{4\lambda[\sigma_\varepsilon^2(1-\beta) + \sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}]}. \tag{48}$$

As one would expect, $g^*$ is a function of $\beta$. The equation for $g^*$ (48) importantly indicates that, under the traditional portfolio problem, the extent to which assets share stochastic trends does indeed matter to investors in terms of portfolio decisions. Unfortunately, $g^*$ (48) does not indicate exactly how the value of $\beta$ relates to the mixture of assets in a portfolio. In order to attempt to determine how investors shift their asset holdings when presented with increasingly related returns (*i.e.*, the case where $\beta$ increases) I derive the partial derivative of $g^*$ (48) with respect to $\beta$, which is given by:

$$\frac{\partial g^*}{\partial \beta} = -\frac{1}{4}\,\frac{\sigma_\varepsilon^2[2\lambda(\sigma_a^2 - \sigma_b^2) - \mu_a + \mu_b]}{\lambda[\sigma_\varepsilon^2(1-\beta) + \sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}]^2}. \tag{49}$$

Perhaps somewhat surprisingly, the partial derivative (49) is impossible to sign in general; however, its ambiguity is actually extremely elucidating. First, consider the denominator. Since the only part that can make the denominator negative is within the squared term, the denominator is always positive.

Next consider the numerator, which can be rearranged into two different, but very insightful formulations:

$$-\sigma_\varepsilon^2[(\mu_b - 2\lambda\sigma_b^2) - (\mu_a - 2\lambda\sigma_a^2)] \tag{50}$$

$$-\sigma_\varepsilon^2[(\mu_b - \mu_a) - 2\lambda(\sigma_b^2 - \sigma_a^2)]. \tag{51}$$

The first formulation illustrates the risk-return tradeoff investors demand: if asset $B_t$ offers the investor greater compensation for risk than asset $A_t$, $g^*$ is a *monotonically* decreasing function of $\beta$ and the investor increases the weight of asset $B_t$ in his portfolio (recall also that $\lambda$ is an indicator of the amount an investor needs to be compensated for increasing risk). Conversely, if $A_t$ offers greater compensation for risk than $B_t$, then $g^*$ is monotonically increasing in $\beta$.

The second formulation provides a slightly more obvious example. If the variances of both $\eta_{at}$ and $\eta_{bt}$ are equal (*i.e.,* $\sigma_a^2 = \sigma_b^2$) then the investor optimizes his holdings by shifting wealth towards the asset with a greater return. Again, the greater the value of $\beta$, the greater impact of a drift in the log prices (1) and (4).

Lastly, there are two very important aspects of the partial derivative of $g^*$ with respect to $\beta$ (49). The first is that $\sigma_\varepsilon^2$ does not affect the sign of the derivative (49). Indeed, since this is common to both, the only important variances are those of the idiosyncratic shocks which indicate which asset is riskier. The second is that it is only the variance of the idiosyncratic (temporary) shocks and the expected returns that affect the optimal portfolio holdings as $\beta$ changes. This is because over one period, regardless of the level of $\beta$, there still exist opportunities for short-term diversification since $\eta_{at}$ is not perfectly (and positively) correlated with $\eta_{bt}$ in general. Thus, it is still optimal for the investor to own both assets.

### *5.2 Numerical examples for the portfolio weight*

Table 1 displays a simple set of numerical examples to illustrate the findings of Section 5.1. To calculate the numerical examples, I assign values to $\mu_a$, $\mu_b$, $\sigma_\varepsilon^2$, $\sigma_a^2$, $\sigma_b^2$, $\sigma_{ab}$, and investor risk tolerance, $\lambda$. French and Poterba (1990) find $\lambda = 3$ to be appropriate and so I use this value throughout the section. The parameters are then held constant while I vary $\beta$ and calculate the optimal portfolio weight on asset $A_t$ (*i.e.,* I calculate $g^*$).

In the first three rows of Table 1, $\mu_a$ is greater than $\mu_b$ and the variance of the stochastic trends is greater than those of the idiosyncratic shocks. I then calculate $g^*$ for the three cases of $\beta = 0, \frac{1}{2}$ and 1. I find that as $\beta$ increases the weight on the home asset, $g^*$ increases also (*i.e.,* as the weight on the shared stochastic trend ($\beta$) increases, the weight on asset $A_t$ also increases). In the last three rows of Table 1, the parameters again stay constant; however, now the idiosyncratic shocks have a greater variance than the stochastic trends. Again, as $\beta$ increases, so too does the weight on the home asset, although to a lesser extent.

**Table 1: Numerical Examples for the Portfolio Weight**

| $\mu_a$ | $\mu_b$ | $\sigma_\varepsilon^2$ | $\sigma_a^2$ | $\sigma_b^2$ | $\sigma_{ab}$ | $\beta$ | $g^*$ |
|---|---|---|---|---|---|---|---|
| 10 | 5 | 5 | 2 | 2 | 1 | 0 | 0.560 |
| 10 | 5 | 5 | 2 | 2 | 1 | $\frac{1}{2}$ | 0.593 |
| 10 | 5 | 5 | 2 | 2 | 1 | 1 | 0.708 |
| 10 | 5 | 5 | 10 | 10 | 1 | 0 | 0.518 |
| 10 | 5 | 5 | 10 | 10 | 1 | $\frac{1}{2}$ | 0.520 |
| 10 | 5 | 5 | 10 | 10 | 1 | 1 | 0.523 |

The results of the numerical example illustrate three issues very clearly. First, the more related the permanent shocks are (*i.e.*, the greater $\beta$) the more investors shift their portfolio holdings toward the asset with the greater compensation for risk. Second, over a short time horizon investors are strongly influenced by short term diversification opportunities as evidenced by the low sensitivity of $g^*$ to changes in $\beta$ when the idiosyncratic shocks have greater variances. Third, the natural reaction from the investor's perspective is therefore to move his portfolio weights toward the most attractive asset, not necessarily the domestic asset. The example in Table 1 illustrates this well because $g^*$ is actually the same for both foreign and domestic investors.

Thus, to develop a model in which home bias can be explained, I must not only determine why a particular country's, say country $A$'s, investors own mostly domestic assets, I must simultaneously determine why citizens of another country, say country $B$, do not similarly own the assets of the country $A$ in large quantities. Fortunately, the theory developed above is able to incorporate situations where I can explain both why people from country $A$ hold mostly asset $A_t$, and why people from country $B$ would hold mostly asset $B_t$. For example, consider a case where either (i) people are overly optimistic about domestic returns, or (ii) transaction costs (including taxes) make the foreign assets more costly to hold than the domestic assets (*i.e.*, lower the mean return). Using either scenario, it is possible that people from country $A$ would have the beliefs indicated in Table 1, but that people from country $B$ would instead believe that $\mu_b = 10$ and $\mu_a = 5$ (*i.e.*, everyone agrees on the variance terms, but they disagree on the mean returns). Table 2 indicates the implications for portfolio weights of such beliefs for those in country $B$. The results of this example certainly typify a home bias for both nations.

**Table 2: Numerical Examples for the Portfolio Weight for Country $B$**

| $\mu_a$ | $\mu_b$ | $\sigma_\varepsilon^2$ | $\sigma_a^2$ | $\sigma_b^2$ | $\sigma_{ab}$ | $\beta$ | $g^*$ |
|---|---|---|---|---|---|---|---|
| 5 | 10 | 5 | 2 | 2 | 1 | 0 | 0.440 |
| 5 | 10 | 5 | 2 | 2 | 1 | $\frac{1}{2}$ | 0.407 |
| 5 | 10 | 5 | 2 | 2 | 1 | 1 | 0.292 |
| 5 | 10 | 5 | 10 | 10 | 1 | 0 | 0.482 |
| 5 | 10 | 5 | 10 | 10 | 1 | $\frac{1}{2}$ | 0.480 |
| 5 | 10 | 5 | 10 | 10 | 1 | 1 | 0.477 |

Notice that in Table 2 as $\beta$ increases, the optimal weight on $A_t$, $g^*$, falls. In other words, as $\beta$ increases, the weight investors from country $B$ place on asset $B_t$ increases as well, indicating that the more related $A_t$ and $B_t$ are, the more pronounced home bias will be in investor portfolios from both $A$ and $B$.

Lastly, it is also conceivable that the perceived risk of foreign markets is greater than that of the domestic market. Table 3 illustrates such a case. Specifically, in the first three rows of Table 3 I set the mean returns equal, but make the variance of country $B$'s idiosyncratic shock greater than that of country $A$'s. In the last three rows of Table 3, I reverse this and make the variance of country $A$'s idiosyncratic shocks greater than those of country $B$'s (*i.e.,* each country views the foreign asset as riskier than the domestic one).

**Table 3: Numerical Examples for the Portfolio Weight when Domestic Asset is Less Risky**

| Country $A$'s beliefs and portfolio weights | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\mu_a$ | $\mu_b$ | $\sigma_\varepsilon^2$ | $\sigma_a^2$ | $\sigma_b^2$ | $\sigma_{ab}$ | $\beta$ | $g^*$ |
| 5 | 5 | 5 | 2 | 4 | 1 | 0 | 0.611 |
| 5 | 5 | 5 | 2 | 4 | 1 | $\frac{1}{2}$ | 0.654 |
| 5 | 5 | 5 | 2 | 4 | 1 | 1 | 0.750 |

| Country $B$'s beliefs and portfolio weights | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\mu_a$ | $\mu_b$ | $\sigma_\varepsilon^2$ | $\sigma_a^2$ | $\sigma_b^2$ | $\sigma_{ab}$ | $\beta$ | $g^*$ |
| 5 | 5 | 5 | 4 | 2 | 1 | 0 | 0.389 |
| 5 | 5 | 5 | 4 | 2 | 1 | $\frac{1}{2}$ | 0.346 |
| 5 | 5 | 5 | 4 | 2 | 1 | 1 | 0.250 |

Similar to the case in Table 1 where the domestic asset in country $A$ has a higher mean perceived return from its citizens' point of view, when the variance of the foreign return is perceived to be greater than that of the home return (*e.g.,* from country $A$'s perspective $\sigma_a < \sigma_b$), there is increasing home bias as $\beta$ increases (*i.e.,* as the correlation between $a_t$ and $b_t$ increases). If the perceptions assumed in Table 3 are reversed in terms of the variances of the idiosyncratic shocks, the portfolio weights chosen by residents of country $B$ are exactly $1 - g^*$. Again recall that a fall in $g^*$ for investors in country $B$ implies lower holdings of country $A$'s asset, so home bias is again more apparent.

In summary, the theory developed in Sections 3 to Section 5 is very easily extended to scenarios where home bias is explained by varying perceptions of risk and mean returns. Moreover, the results in this section indicate that, in the presence of cointegrated markets, investors need not be nearly as overly optimistic about home market returns as French and Poterba (1991) concluded (*i.e.,* home bias is not necessary as puzzling as previously thought).

### *5.3 Exponential utility and optimal portfolio allocations*

While the results from Section 5.1 are clear and indicate that $g^*$ is indeed a function of $\beta$, I also use an alternative objective function. For their important study of portfolio home bias, French and Poterba (1991) employed an exponential Constant Absolute Risk Aversion (CARA) utility function. Specifically, they used an exponential utility function defined over current wealth, $w$, and initial wealth, $w_0$, and given by: $U(w) = -\exp(\frac{-\lambda w}{w_0})$. In the context of a portfolio of two risky assets (like the one considered in Section 5.1), with a vector of weights $g$, a vector of means $\mu$, and a variance-covariance matrix of the returns $\Sigma$, the expected utility function is given by:

$$E[U(W)] = -\exp\left[-\lambda\left(g\mu - \frac{\lambda g'\Sigma g}{2}\right)\right]. \tag{52}$$

Expected utility (52) can be rewritten in scalar form, however, by using the expectation and variance of the portfolio (46), for the mean vector $\mu$, the variance-covariance matrix $\Sigma$, and the weighting vectors respectively. This rearrangement gives:

$$E[U(W)] = -\exp\Bigg[-\lambda\Bigg(g\mu_a + (1-g)\mu_b \\ - \frac{\lambda[g^2(\sigma_\varepsilon^2 + 2\sigma_a^2) + (1-g)^2[\sigma_\varepsilon^2 + \sigma_b^2] + 2g(1-g)(\beta\sigma_\varepsilon^2 + \sigma_{ab})]}{2}\Bigg)\Bigg]. \tag{53}$$

As with the traditional portfolio problem, the objective is to maximize $E[U(W)]$ by selecting the optimal portfolio weights, $g^*$. To do so, I similarly differentiate equation (53) with respect to $g$ which simplifies to:

$$\frac{\partial E[U(W)]}{\partial g} = \mu_a - \mu_b - \frac{\lambda[2g(\sigma_\varepsilon^2 + 2\sigma_a^2) - 2(1-g)[\sigma_\varepsilon^2 + 2\sigma_b^2] + (2-4g)(\beta\sigma_\varepsilon^2 + \sigma_{ab})]}{2}. \tag{54}$$

The first order condition (54) is remarkably similar to the first order condition for the traditional portfolio problem (47). Accordingly, equating the first order condition to zero and rearranging to solve for $g^*$ gives nearly the same result as before (48):

$$g^* = \frac{\mu_a - \mu_b + 2\lambda[\sigma_\varepsilon^2(1-\beta) + \sigma_b^2 - 2\sigma_{ab}]}{\lambda[2\sigma_\varepsilon^2(1-\beta) + \sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}]}. \tag{55}$$

As one would expect, the partial derivative of equation (55) is extremely similar to that of equation (48) and given by:

$$\frac{\partial g}{\partial \beta} = -\frac{1}{2}\,\frac{\sigma_\varepsilon^2[\lambda(\sigma_a^2 - \sigma_b^2) - \mu_a + \mu_b]}{\lambda[\sigma_\varepsilon^2(1-\beta) + \sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}]^2}. \tag{56}$$

The interpretations for the partial derivatives of $g^*$ with respect to $\beta$ for both utility functions (equations (56) and (49)) are the same and so I do not provide them again (see Section 5.1 for discussion).

### *5.4 Multi-period returns and the traditional portfolio problem*

In Section 4.2 I found that as the investment horizon increases, the correlation between the returns also increases, so long as $\beta \neq 0$. It seems reasonable, therefore, to expect that as the investment horizon increases so too does the optimal weight, $g^*$. To examine this hypothesis, I use the same steps as in Section 5.1, but use the returns defined in equations (38) and (39), instead. Maximizing the objective function and rearranging for $g^*$ now gives:

$$g^* = \frac{n(\mu_a - \mu_b) + 2\lambda[n\sigma_\varepsilon^2(1-\beta) + 2\sigma_b^2 - 2\sigma_{ab}]}{4\lambda[n\sigma_\varepsilon^2(1-\beta) + \sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}]}. \tag{57}$$

where $n$ is again the time horizon of the investor. Dividing both the numerator and denominator of equation (4.2) by a factor of $n$ gives:

$$g^* = \frac{\mu_a - \mu_b + 2\lambda\left[\sigma_\varepsilon^2(1-\beta) + \frac{2\sigma_b^2 - 2\sigma_{ab}}{n}\right]}{4\lambda\left[\sigma_\varepsilon^2(1-\beta) + \frac{\sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}}{n}\right]}. \tag{58}$$

Finally, taking the limit of $g^*$ (58) as $n$ tends to infinity gives:

$$\lim_{n\to\infty} g^* = \frac{\mu_a - \mu_b}{4\lambda\sigma_\varepsilon^2(1-\beta)} + \frac{1}{2}. \tag{59}$$

If $\beta = 1$ in equation (59), $g^*$ is driven to either positive or negative infinity; however, because the investor is limited to $g \in [0, 1]$, the limit of $g^*$ (59) really indicates that the investor will choose either $g^* = 1$ or $g^* = 0$ depending which asset offers the greatest possible return. Importantly, the limit of $g^*$ (59) also indicates that the original hypothesis was correct: as the time horizon tends to infinity the investor will be driven to a corner solution. In other words, over a long time horizon, when short term gains from diversification will not be realized (precisely because they are short term) the investor will invest in which ever asset has the greatest expected return.

### *5.5 Optimal portfolios with weights chosen jointly with consumption*

The theoretical analysis in Sections 5.1 to 5.4 was all carried out without regard for the lifetime utility maximization problem facing every investor. Instead, it was implicitly assumed that each investor tried to maximize terminal wealth. As a final robustness check, I also explore the long-term portfolio choices a theoretical investor would make while jointly maximizing his lifetime utility.

To do so, I draw on the work by Samuelson (1969). As Danthine and Donaldson (2005) describe the problem, the investor chooses consumption, $c_t$, and portfolio weights, $g_t$, to maximize his lifetime utility over consumption, using a discount factor $\delta$. The constraint is that his current wealth, $w_t$, is equal to his labour earnings, $y_t$, minus his consumption plus his earnings from his portfolio. Specifically, the problem can be represented as:

$$\max_{c} \quad E_t \left[ \sum_{i=0}^{\infty} \left( \frac{1}{1+\delta} \right)^i U(c_{t+i}) \right] \tag{60}$$

$$s.t. \quad w_t = \left[ 1 + g_t r_{a,t-1} + (1 - g_t) r_{b,t-1} \right] w_{t-1} + y_t - c_t \tag{61}$$

$$w_0 \ \text{ given.} \tag{62}$$

While more realistic, the above formulation can only be solved explicitly if the utility function is of the constant relative risk aversion family, and labour income is constant. Otherwise, numerical approximations must be used.

The crucial point, however, is that, with the above objective function and constraints, the optimal portfolio weights — much like the optimal weights found in Sections 5.1 and 5.3 — depend on the expected returns and the correlation between them. While most formulations include a risk free asset, there is no reason to think that the results would not be similar for a problem without one. Moreover, a formal derivation of the optimal weights given the above problem is beyond the scope of this essay. My purpose here instead is to acknowledge this important area of the literature and indicate that including a consumption objective should not change the fundamental result. Indeed, the crucial point is still that cointegration of national markets reduces the gains from diversification, thus reducing the cost of a home bias to investors. Now, having developed a theoretical framework in which cointegrated markets could, at least partially, explain portfolio home bias I conduct an empirical examination of the G7 markets to test for cointegration.

## 6 Countries and Data Sources

### *6.1 Data source*

A scarcity of data is, of course, the largest obstacle in terms of countries to include and time period to study. In this respect, I do my best to follow the sources used by Chang (2001), Allen and MacDonald (1995), Kasa (1992), and others, and utilize the *total return indices* developed and maintained by Morgan Stanley Capital International Barra (MSCI hereafter). Like other authors, I include dividends since they are extremely important to long-term investors and they are also treated very differently across countries (and are in part affected by the tax code of the country where the stock is listed). While there are many alternatives to the MSCI indices (*e.g.,* the total return variants of the Dow Jones Industrial Average on the NYSE and the S&P/TSX Composite on the Toronto Stock Exchange) the MSCI indices have the benefit of being constructed by the same firm with the focus on reflecting the same asset mixes. Moreover, Kanas (1998) notes that MSCI also use a very large base of firms to construct its indices, further reducing the likelihood of individual sectors dominating the relationships. Research by Roll (1992) indicates that heterogeneity in index construction can cause substantial divergence in behavior. In other words, since the indices are mostly homogenous across stock markets, they are more appropriate for detecting *national* as opposed to *sectoral* trends.

The MSCI total return indices have two variants for each country: an index incorporating *net* dividends and an index incorporating *gross* dividends. The net dividend indices aim to reflect the *minimum* possible dividend reinvestment by incorporating withholding taxes. Alternatively, the gross dividend indices aim to reflect the value of the index under the maximum possible dividend reinvestment. For exact definitions, see MSCI (2009). In determining if there exists international diversification opportunities for investors, it seems that the *net* indices represent the worst case scenario while the *gross* indices represent the best case scenario. Thus, since both types of indices appear desirable in certain contexts, and using both indices would act as a robustness test, I have decided to use both types of indices throughout the empirical section.

### *6.2 Countries examined*

With the data limitations in mind, I had to restrict my analysis to more “developed” markets. Moreover, given Canada’s position in the G7 it seems reasonable to hypothesize that there may be higher levels of integration between Canada’s equity markets and those of other member nations than the level of integration between Canada’s equity markets and non-member nations. Simultaneously, it is also likely that nations with strong export/import ties to Canada may have similarly strong ties to Canada’s equity markets.

Therefore, I have decided to examine the possibility of cointegrated stock markets primarily between Canada and the G7 nations, which are Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States. Having selected the countries to include, I first need to take the natural logarithm of the level of the index, since the theoretical sections above have focused on this transformation of the indices. Figures 1 and 2 graph the weekly log levels of the net and gross total return indices for Canada, Japan, Italy, and the United States for the period of 1999-2009. Table 4 presents the summary statistics for the logs of the total return indices for the countries selected.

**Figure 1: Net Indices: 1999-2009**

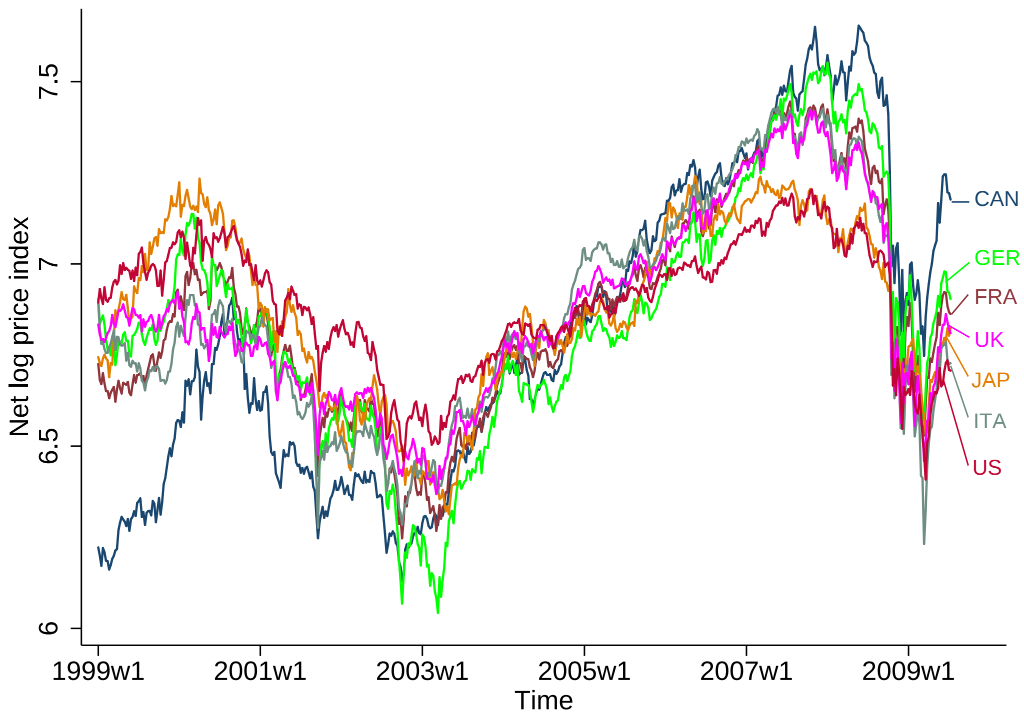
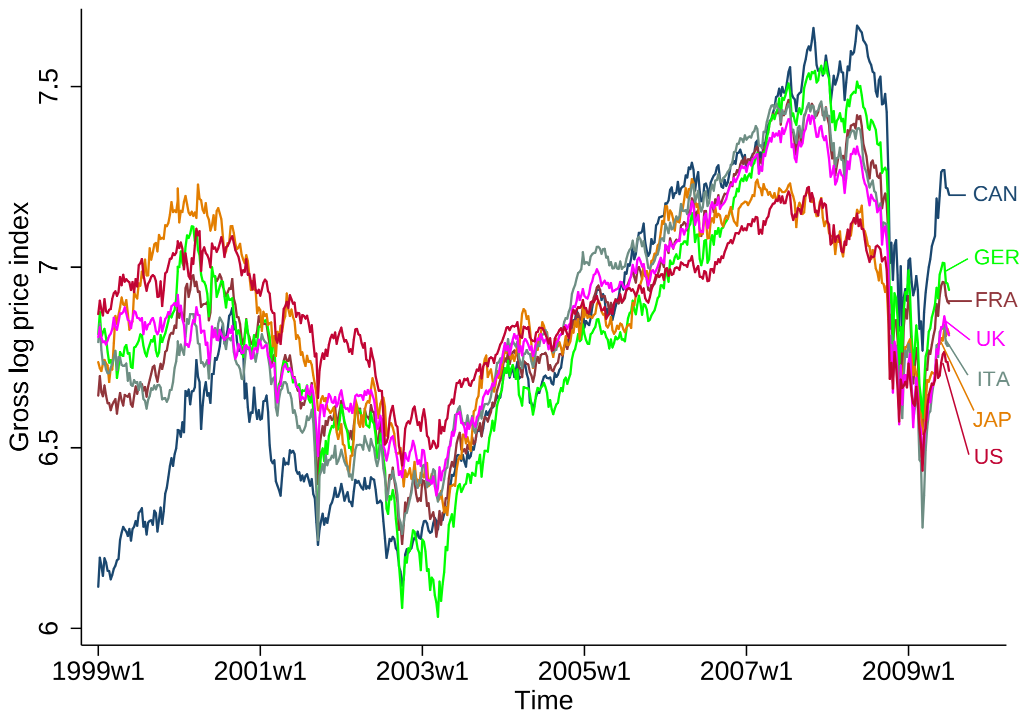



**Figure 2: Gross Indices: 1999-2009**




Both series from weekly data and normalized so each index has same mean; Source: Bloomberg

**Table 4: Summary Statistics for the G7 Stock Indices, 1999-2009**

| Country | Mean | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|
| Net Total Return Indices | | | | |
| Canada | 7.735 | 0.4253 | 7.05 | 8.573 |
| France | 8.106 | 0.2935 | 7.489 | 8.687 |
| Germany | 7.903 | 0.3379 | 7.095 | 8.605 |
| Japan | 8.218 | 0.2432 | 7.652 | 8.58 |
| Italy | 6.689 | 0.3025 | 6.06 | 7.259 |
| United Kingdom | 8.156 | 0.264 | 7.653 | 8.705 |
| United States | 7.848 | 0.1751 | 7.364 | 8.159 |
| Gross Total Return Indices | | | | |
| Canada | 7.954 | 0.4375 | 7.26 | 8.814 |
| France | 8.428 | 0.3042 | 7.801 | 9.031 |
| Germany | 8.179 | 0.3459 | 7.364 | 8.9 |
| Japan | 8.319 | 0.2436 | 7.753 | 8.684 |
| Italy | 6.973 | 0.3203 | 6.362 | 7.576 |
| United Kingdom | 8.448 | 0.264 | 7.944 | 8.997 |
| United States | 8.198 | 0.1771 | 7.743 | 8.528 |

Data is weekly and in logged levels (from 2 January 1999 to 3 July 2009).
Source: Bloomberg

## 7 Unit Root Specification

In Section 3.1, I noted that while there may be an epistemological debate in econometrics regarding the validity of the random walk model of stock prices, I do not add to that literature. Instead, I utilize the current unit root tests and interpret them as either evidence for or against the random walk hypothesis, while bearing in mind the well known issues.

The most widely accepted test for unit roots is the augmented Dickey and Fuller

(1979) test (ADF test hereafter). In particular, the test is given by the following:

$$\Delta y_{xt} = \alpha + \rho y_{x,t-1} + \gamma t + \sum_{j=1}^{p} \theta_j \Delta y_{x,t-j} + \vartheta_t \tag{63}$$

where $\gamma = 0$ in an ADF test conducted with no linear trend term, $p$ is the number of lags which, in this case, is 1, $y_{xt}$ is a scalar, and $x$ is the country in the regression.

The null hypothesis is that there is a unit root (*i.e.*, $\rho = 0$ in equation (63)), and, therefore that the variable $y_{xt}$ is integrated of order one, or $I(1)$. Given the well known size and power problems associated with ADF tests, I use the 1% critical value as my rejection rule.

Lag selection is very important for testing for cointegration. Thus, for each series, before I conduct the ADF tests, I use Schwarz's Bayesian information criterion (SBIC hereafter) — rather than Akaike's information criterion since the SBIC is more parsimonious — to select the appropriate number of lags ($p$) to include in the regression. For all series, both net and gross, the SBIC indicated that one lag was optimal. Moreover, including 1 lag appropriately controls for serial autocorrelation in the results in Section 8.4 with respect to the Johansen (1988) pairwise tests.

The presence or absence of a trend term is also important. Thus, to determine if one should be included, I use the first differences of each series and then conduct a $t$-test on the mean. Surprisingly, the test indicates that there is no trend term for all the series. As a result, the appropriate specification for the ADF test is one lag ($p = 1$) and no trend term ($\gamma = 0$).

Table 5 reports the results from ADF tests on both the net and gross indices under the "Levels" heading. The results indicate that all variables have a unit root. Indeed, in absolute terms the largest statistics are both for Germany and are only 1.562 and 1.652 compared to the 1% critical value of 3.960. In other words, there is strong evidence in favour of the unit root hypothesis. To ensure that the data are indeed integrated of order one, however, I also first difference the data and then conduct the ADF test with zero lags (as indicated by the SBIC) and no trend term. Table 5 reports the full results under the "First Difference" heading. As expected, I am easily able to reject the null hypothesis of a unit root for each series at the 1% level. For robustness, I also use the Phillips and Perron (1988) test for both the data in levels and in first differences and again find strong evidence that all the series are integrated of order one. In the interest of brevity, the results of the Phillips and Perron (1988) tests are not reported.

**Table 5: ADF test results**

| | Levels | | | | First Differences | |
|---|---|---|---|---|---|---|
| | No Trend | | Trend | | No Trend | |
| Country | Net | Gross | Net | Gross | Net | Gross |
| Canada | -1.304 | -1.418 | -1.456 | -1.46 | -25.36 | -25.53 |
| France | -1.381 | -1.391 | -1.381 | -1.425 | -24.05 | -24.16 |
| Germany | -1.308 | -1.27 | -1.562 | -1.652 | -22.98 | -23.1 |
| Japan | -1.488 | -1.49 | -1.346 | -1.338 | -24.52 | -24.56 |
| Italy | -1.298 | -1.26 | -1.314 | -1.447 | -23.25 | -23.35 |
| United Kingdom | -1.366 | -1.375 | -1.247 | -1.279 | -16.93$^{\dagger}$ | -16.94$^{\dagger}$ |
| United States | -1.711 | -1.705 | -1.53 | -1.493 | -25.34 | -25.34 |
| Critical values | | | | | 1% | 5% |
| No Trend | | | | | -3.960 | -3.410 |
| Trend | | | | | -3.430 | -2.860 |

Data is weekly and in logged levels (from 2 January 1999 to 3 July 2009).
Source: Bloomberg
$^{\dagger}$ ADF included one lag

Since I have now found that all variables are integrated of the same order (necessary for the Engle and Granger (1987) two-step procedure and the Zivot and Andrews (1992) tests, but not necessary for the Johansen (1988) procedure), it seems highly appropriate to use a cointegration framework and determine if there is evidence that portfolio home bias may simply be an optimal response from investors. Thus, in Section 8 I employ pairwise tests for cointegration and in Section 9 I use the Johansen (1988) procedure to test for cointegration in the system as a whole. Recall that if national indices are cointegrated (*i.e.,* $\beta = 1$), the implication is that investors have fewer opportunities for diversification. In other words, a home bias in investors' portfolios is less "puzzling" if test results indicate that there are cointegrating vectors within the G7 nations.

## 8 Testing for Pairwise Cointegration

### *8.1 Why a pairwise approach?*

While international investors will most likely consider investing in multiple countries at one time, I test for cointegration using a pairwise approach for three main reasons: (i) the systems approach may not find all cointegrating vectors in a finite sample (due to the size and power issues of the Johansen (1988) procedure) and so the gains from diversification may be exaggerated, (ii) pairwise cointegration still has important economic implications — if there are pairs of countries with cointegrated indices, investors from either country should diversify their holdings elsewhere (*e.g.,* if the Canadian and US markets are cointegrated, it would not make sense for long-term Canadian investors to buy US assets for the purpose of diversification) — and (iii) there are many more testing procedures available if a pairwise approach is used, so robustness checks are easier under a bivariate specification. As Allen and MacDonald (1995) note, a pairwise analysis is instructive since it "demonstrates which series are moving together in the long-run." Moreover, many studies have used the bivariate approach, for examples see Kanas (1998), Allen and MacDonald (1995), and Chang (2001). Thus, in Sections 8.2 to 8.4 I conduct pairwise analysis, while in Section 9 I use a systems approach.

### *8.2 Engle and Granger (1987) two-step*

Although there are many testing procedures for cointegration, one of the simplest and oldest tests is the Engle-Granger two-step. Specifically, if two variables are both I(1) and cointegrated and one is regressed on another, the residuals from the regression will be stationary (or I(0)). To carry out this test, I follow the method proposed by Engle and Granger (1987) as explained by Enders (1995). Specifically, I run the following regressions:

$$y_{xt} = \alpha_{0xz} + \alpha_{1xz} y_{zt} + \epsilon_{xz,t}, \qquad \forall\ x \neq z \tag{64}$$

$$\Delta\hat{\epsilon}_{xz,t} = \rho\hat{\epsilon}_{xz,t-1} + \sum_{j=1}^{p} \theta_j \Delta\hat{\epsilon}_{xz,t-j} + e_t \tag{65}$$

where $y_{xt}$ and $y_{zt}$ are the logged levels of the total return indices (gross or net) in countries $x$ and $z$ respectively.

In the first step, equation (64), I regress the index level of one country onto that of another and save the residuals. In the second step, equation (65), I perform an ADF test with no constant (since the regressand is a residual none is needed) but 2 lags since there is autocorrelation in the errors and two lags was selected by the Bayesian information criterion. If the two variables are cointegrated (*i.e.,* the ADF test indicates that in equation (65) $\rho \leq 1$) the slope coefficient, $\alpha_{1xz}$, can be interpreted as the 'strength' of the cointegrating relation since it measures the sensitivities — the change in the index level in the regressand country brought about by a one percentage point increase in index level of the regressor country. I report the full results in Tables 6 and 7 below.

**Table 6: Engle and Granger (1987) ADF tests — Net Indices**

| | France | Germany | Japan | Italy | United Kingdom | United States |
|---|---|---|---|---|---|---|
| Canada | -1.919 | -2.661 | -1.194 | -1.886 | -1.804 | -0.9936 |
| France | | -2.778 | -2.393 | -1.911 | -2.503 | -1.809 |
| Germany | | | -2.61 | -1.09 | -1.961 | -1.514 |
| Japan | | | | -2.202 | -2.339 | -3.285* |
| Italy | | | | | -3.701** | -1.824 |
| United Kingdom | | | | | | -1.769 |
| Critical values | | | | 1% | 5% | 10% |
| | | | | -3.430 | -2.860 | -2.570 |

Data is weekly and in logged levels (from 2 January 1999 to 3 July 2009).
Source: Bloomberg
** denotes significance at 1% level; * denotes significance at 5% level

The tables indicate that there are very few long-run relationships between the equity markets of the G7 countries. Indeed, the only country pairs with cointegrated indices are Italy and the United Kingdom and Japan and the United States. Importantly, it does not appear that the use of either gross or net indices is affecting the results in this case since the results are remarkably consistent across both.

**Table 7: Engle and Granger (1987) ADF tests — Gross Indices**

| | France | Germany | Japan | Italy | United Kingdom | United States |
|---|---|---|---|---|---|---|
| Canada | -1.998 | -2.511 | -1.009 | -2.194 | -1.683 | -0.7179 |
| France | | -2.63 | -1.978 | -2.073 | -2.335 | -1.283 |
| Germany | | | -2.321 | -1.164 | -1.795 | -1.088 |
| Japan | | | | -2.005 | -2.192 | -3.463** |
| Italy | | | | | -3.403** | -1.543 |
| United Kingdom | | | | | | -1.436 |
| Critical values | | | | 1% | 5% | 10% |
| | | | | -3.430 | -2.860 | -2.570 |

Data is weekly and in logged levels (from 2 January 1999 to 3 July 2009).
Source: Bloomberg
** denotes significance at 1% level; * denotes significance at 5% level

### *8.3 Zivot and Andrews (1992) tests*

While the Engle and Granger (1987) two-step procedure provides a quick way and relatively simple method of determining if two variables are cointegrated, it has limitations. Specifically, if there is a *structural break* in either the trend or the intercept, the Engle and Granger (1987) procedure may yield incorrect results. To allow for this possibility I employ Zivot and Andrews (1992) tests for cointegration which allow for an *unknown* structural break in the relationship (in either the trend term or the intercept). A discussion of how this test works is beyond the scope of this paper; however, interested readers should see Maddala and Kim (1998, 401) and Zivot and Andrews (1992).

The Zivot and Andrews (1992) tests are conducted in a similar fashion as the Engle and Granger (1987) test for cointegration: I first regress the logged level of each country's total return index (gross and net) on one another using equation (64) and save the residuals (*i.e.,* $\alpha_{xz}$ has the same interpretation as in Section 8.2). I then use the Zivot and Andrews (1992) test (with two lags, similar to the Engle and Granger (1987) two-step above) to test if the residuals are random walks. Like the ADF test, the Zivot and Andrews (1992) test has a null hypothesis of a unit root in the data (*i.e.,* the null hypothesis is that there is no cointegration). Tables 8 and 9 report the full results below.

As with the Engle and Granger (1987) tests, the results of the Zivot and Andrews

**Table 8: Zivot and Andrews (1992) tests — Net Indices**

| | France | Germany | Japan | Italy | United Kingdom | United States |
|---|---|---|---|---|---|---|
| Canada | -3.682 | -3.934 | -4.392 | -4.406 | -3.087 | -4.949 |
| France | | -3.972 | -4.849 | -4.545 | -5.437* (2000w38) | -8.924** (2001w17 ) |
| Germany | | | -5.117* ( 2007w13) | -4.238 | -4.016 | -6.149** (2005w20 ) |
| Japan | | | | -4.793 | -4.117 | -4.297 |
| Italy | | | | | -5.074 | -5.379* (2002w24) |
| UK | | | | | | -4.991 |
| Critical values | | | | | 1% | 5% |
| | | | | | -5.57 | -5.08 |

Data is weekly and in logged levels (from 2 January 1999 to 3 July 2009). Source: Bloomberg
** denotes significance at 1% level; * denotes significance at 5% level
Date of structural break indicated by year and week in parenthesis (if significant)

**Table 9: Zivot and Andrews (1992) tests — Gross Indices**

| | France | Germany | Japan | Italy | United Kingdom | United States |
|---|---|---|---|---|---|---|
| Canada | -4.696 | -4.542 | -4.236 | -5.094* (2006w33) | -3.442 | -4.579 |
| France | | -3.85 | -4.732 | -4.765 | -4.932 | -6.474** (2003w8) |
| Germany | | | -4.618 | -4.338 | -3.921 | -5.688** (2006w33) |
| Japan | | | | -5.05 | -4.242 | -4.285 |
| Italy | | | | | -5.138* (2002w43) | -6.426** (2002w14) |
| UK | | | | | | -5.025 |
| Critical values | | | | | 1% | 5% |
| | | | | | -5.57 | -5.08 |

Data is weekly and in logged levels (from 2 January 1999 to 3 July 2009). Source: Bloomberg
** denotes significance at 1% level; * denotes significance at 5% level
Date of structural break indicated by year and week in parenthesis (if significant)

(1992) are relatively consistent between both net and gross indices. The results indicate however, that once regime shifts are taken into account, the markets of the G7 are strongly related to that of the United States — much more so than the results in Section 8.2. Given the fundamental changes that have taken place over the past decade in the world of finance it should not be particularly surprising that structural breaks do matter in this case. Indeed, for both index types, the number of cointegrating vectors increases from 2 in the Engle and Granger (1987) framework to 5 in the Zivot and Andrews (1992) framework. Moreover, most of these additional vectors are from regressions including the US market. Specifically, France, Germany and Italy are cointegrated with the US indices, often at the 1% confidence level, after I allow for structural breaks. Notice too that the breaks were scattered throughout the years between 2001 and 2007 and only one of the breaks in 2001 itself (France). Also, it is important to note that the Zivot and Andrews (1992) framework is unable to examine structural breaks at the tails of the observation periods. Specifically, I could not explore the possibility of a structural break in the first or last 15% of observations — thus eliminating all of 2009 and much of 2000 and 2008 from the years that were allowed to have a structural break.

### *8.4 Johansen (1988) test procedure*

The Engle and Granger (1987) and Zivot and Andrews (1992) procedures I used in Section 8 both utilize the residuals of OLS regressions to determine if the two variables are cointegrated. The Johansen (1988) procedure, on the other hand, estimates the number of cointegrating vectors directly from the vector autoregression (VAR) representation itself:

$$Y_t = \mu + \gamma t + u_t. \tag{66}$$

For a complete derivation see Johansen (1988) or Juselius and Hendry (2000). Using this procedure, only one step is required; however, much more about the structure of the data itself must be considered.

Under the Johansen (1988) procedure, it is imperative to properly control for deterministic terms. To do so, I again use the results of the $t$-tests on the means of the first differences of the logged levels of each series from Section 8.2. Since each series, when first differenced, has a mean of zero, I elect to restrict my model to one having no trend term as suggested by Juselius and Hendry (2000) and an unrestricted constant term. Having now chosen the appropriate deterministic terms I can utilize the Unobserved Components Representation — slightly different than the Vector Error Correction Model (VECM) — of the system as follows:

$$\Delta Y_t = \Pi(Y_{t-1} - \gamma(t-1) - \mu) + \mu^* + \sum_{j=1}^{p-1} c_j \Delta Y_{t-j} + \varepsilon_t \tag{67}$$

$$Y_t = [y_{xt} \;\; y_{zt}]' \qquad \forall\, x \neq z$$

where the number of cointegrating vectors ($r$) is equal to $rank(\Pi)$, and $p$ is the number of lags. I then use this representation to study the possibility of pairwise cointegration within the G7.

The Johansen (1988) method in this case tests the null hypothesis, $H_0 : r \leq q$ against the alternative, $H_A : r > q, \;\; q = 0, 1$. For example, suppose the test were to reject $H_0 : r = 0$, but not reject $H_0 : r = 1$. This indicates that there is a cointegrating vector. Of course, because the series all have unit roots, there can be no more than one cointegrating vector for each pairwise test. Note also that I follow the usual practice and use $r$ to indicate the number of cointegrating vectors — not be confused with the notation used for returns, $r_{xt}$, which always has a time $t$ and country $x$ index in the subscript.

Lastly, since there are well known power issues with respect to the Johansen (1988) method (*i.e.*, the test does not reject a *false* null hypothesis often enough in finite samples), I use the Reimers (1992) finite sample correction, and 1% critical values (see Cheung and Lai (1993) for a discussion). Specifically, I multiply each statistic by:

$$\frac{T - pm}{T} \tag{68}$$

where $T$ is the number of time observations, $p$ is the number of lags, and $m$ is the number of left hand variables. In this case, the correction is: $\frac{547-(1)(2)}{547} \simeq 0.9963$. I report the results of the Johansen (1988) tests in Tables 10 and 11.

In the interest of brevity, I have only reported the trace statistic and excluded the somewhat less reliable $\lambda_{max}$ statistic. Unlike the results from Section 8, the results of

**Table 10: Johansen (1988) Trace Statistic Results† — Net Indices**

| | France | Germany | Japan | Italy | United Kingdom | United States |
|---|---|---|---|---|---|---|
| Canada | | | | | | |
| $r \leq 0$ | 7.400 | 9.068 | 5.962 | 8.133 | 7.067 | 6.989 |
| $r \leq 1$ | 1.606 | 1.648 | 1.486 | 1.444 | 1.590 | 1.701 |
| France | | | | | | |
| $r \leq 0$ | | 16.011 | 17.263 | 9.167 | 11.372 | 21.393* |
| $r \leq 1$ | | 3.004 | 1.191 | 1.941 | 2.131 | 1.733 |
| Germany | | | | | | |
| $r \leq 0$ | | | 16.957 | 9.661 | 14.124 | 18.136 |
| $r \leq 1$ | | | 1.011 | 4.112 | 3.737 | 1.384 |
| Japan | | | | | | |
| $r \leq 0$ | | | | 16.295 | 19.167 | 20.389* |
| $r \leq 1$ | | | | 1.294 | 1.372 | 1.785 |
| Italy | | | | | | |
| $r \leq 0$ | | | | | 18.235 | 22.807* |
| $r \leq 1$ | | | | | 1.649 | 1.988 |
| United Kingdom | | | | | | |
| $r \leq 0$ | | | | | | 27.818* |
| $r \leq 1$ | | | | | | 2.049 |

Data is weekly and in logged levels (from 2 January 1999 to 3 July 2009). Source: Bloomberg
** denotes significance at 1% level; * denotes significance at 5% level
1% critical value for $H_0 : r \leq 0$ is 20.04
1% critical value for $H_0 : r \leq 1$ is 6.65
†The Reimers (1992) finite sample correction has been applied to all statistics

the Johansen (1988) procedure are *not* consistent between the two types of total return indices used (*i.e.,* the results from the net indices differ significantly from those of the gross indices). In the net indices, I find that there are 4 cointegrating vectors, all of which include the United States in the cointegration relationship. Specifically, I find that the net indices of France, Japan, Italy, and the United Kingdom are all cointegrated with the net index of the United States. On the other hand, when using the gross indices, I find only one cointegrating relationship, between Japan and the United States. Interestingly, the results from the Johansen (1988) tests using net indices look remarkably similar to those of the Zivot and Andrews (1992) results; however, the results from the Johansen (1988) tests using gross indices are extremely similar to those of the Engle and Granger (1987) two-step. It is important to bear in mind that the Johansen (1988) procedure, like the Engle and Granger

**Table 11: Johansen (1988) Trace Statistic Results† — Gross Indices**

| | France | Germany | Japan | Italy | United Kingdom | United States |
|---|---|---|---|---|---|---|
| Canada | | | | | | |
| $r \leq 0$ | 8.415 | 9.385 | 5.285 | 10.006 | 7.776 | 5.014 |
| $r \leq 1$ | 1.776 | 1.706 | 2.012 | 1.472 | 1.711 | 1.799 |
| France | | | | | | |
| $r \leq 0$ | | 15.063 | 8.980 | 9.324 | 11.439 | 6.427 |
| $r \leq 1$ | | 2.882 | 1.784 | 1.810 | 2.152 | 1.949 |
| Germany | | | | | | |
| $r \leq 0$ | | | 10.220 | 9.378 | 13.501 | 8.541 |
| $r \leq 1$ | | | 1.689 | 3.867 | 3.708 | 2.364 |
| Japan | | | | | | |
| $r \leq 0$ | | | | 6.925 | 9.384 | 22.328* |
| $r \leq 1$ | | | | 1.453 | 1.622 | 1.783 |
| Italy | | | | | | |
| $r \leq 0$ | | | | | 15.452 | 6.267 |
| $r \leq 1$ | | | | | 1.574 | 1.541 |
| United Kingdom | | | | | | |
| $r \leq 0$ | | | | | | 7.564 |
| $r \leq 1$ | | | | | | 1.896 |

Data is weekly and in logged levels (from 2 January 1999 to 3 July 2009). Source: Bloomberg
** denotes significance at 1% level; * denotes significance at 5% level
1% critical value for $H_0 : r \leq 0$ is 20.04
1% critical value for $H_0 : r \leq 1$ is 6.65
†The Reimers (1992) finite sample correction has been applied to all statistics

(1987) two-step, does *not* allow for a structural break in the data.

The Johansen (1988) framework is easily extended to a systems approach where the researcher can test if there are cointegrating relationships between more than two variables. With respect to the UCR (67), in a systems framework with $m$ variables $\Pi$ can have $rank(\Pi) \leq m - 1$. In other words there can be up to $m - 1$ cointegrating vectors. While other researchers (*e.g.,* Kasa (1992) and Allen and MacDonald (1995)) have carried out such multivariate examinations, it is not immediately clear exactly what a cointegrating vector involving multiple countries would imply for long-term investors and the possible presence or absence of gains from international diversification. Moreover, since my theoretical exploration was for bivariate cointegration, it seems most appropriate to maintain the pairwise approach in my empirical work.

## 9 Systems Approach

The empirical analysis I have done so far has focused on bivariate tests for cointegration. While a pairwise approach has very important economic implications (as well as intuition), it is somewhat removed from the reality of investors. Indeed, one could use matrix algebra to model seven or more indices as a system of possibly cointegrated variables. I do not derive a multivariate model here; however, because the algebra for such an extension is complex and the economic conclusion is similar to the bivariate case: the fewer stochastic trends there are, the less are the gains from international portfolio diversification, and the less puzzling the evidence for investor home bias.

Thus, as a last application of the theory developed in the first half of this essay, I test for cointegration between any of the G7 indices simultaneously. To do so, I again use the Johansen (1988) procedure from Section 8.4. Indeed, consider the unobserved components representation (67) once more. In the current multivariate framework $Y_t$ becomes:

$$Y_t = [r_{CAN,t} \quad r_{FRA,t} \quad r_{GER,t} \quad r_{JAP,t} \quad r_{ITA,t} \quad r_{UK,t} \quad r_{US,t}]'. \tag{69}$$

As in the bivariate case, $rank(\Pi)$ indicates the number of cointegrating vectors. Similar to the pairwise analysis, the maximum possible number of cointegrating vectors is one less than the number of variables in $Y_t$, which in the case of the G7 stock markets would be six. In a multivariate framework, it is not as clear what an intermediate number of cointegrating vectors (*i.e.,* $5 \geq rank(\Pi) \geq 1$) indicates in terms of long term diversification opportunities for investors. Granger's Representation Theorem (GRT) provides some insights though. In particular, GRT states that the number of variables is equal to the number of cointegrating vectors plus the number of stochastic trends. For example, if there are 5 cointegrating vectors then there would be 2 stochastic trends. In other words, all variables would be some linear combination of the two stochastic trends and an $I(0)$ noise term. More importantly for this essay, however, is that the greater the number of stochastic trends, the lower the less correlated the variables are in the long-run (*i.e.,* the more stochastic trends, the greater the possible gains from diversification).

The deterministic components I found to be appropriate for the pairwise approach

— an unrestricted constant with no trend — are still valid since all of the series had the same properties in this respect. The number of lags included, however, had to be increased. While for both the net and gross indices the SBIC indicated 2 lags were required, serial autocorrelation was present in the residuals when the model was estimated. The higher number of lags required for the system as a whole compared to the pairwise estimates could be caused by a plethora of phenomena. For example, it may be due to one or more volatile exchange rates because all the series are in USD terms (it may well have been a country with no bivariate cointegrating vectors) or a structural break. Nonetheless, in an effort to control for the serial autocorrelation I increased the number of lags to 3, which removed all autocorrelation in the errors. Table 12 reports the results of the Johansen (1988) systems procedure.

**Table 12: Johansen (1988) System Trace Statistic Results**

| Null Hypothesis | 5% Critical Value | Net Indices Trace Statistic† | Gross Indices Trace Statistic† |
|---|---|---|---|
| $r \leq 0$ | 124.24 | 120.17 | 118.55 |
| $r \leq 1$ | 94.15 | 74.85 | 77.01 |
| $r \leq 2$ | 68.52 | 47.7 | 48.35 |
| $r \leq 3$ | 47.21 | 29.3 | 28.46 |
| $r \leq 4$ | 29.68 | 14.47 | 12.59 |
| $r \leq 5$ | 15.41 | 6.18 | 5.22 |
| $r \leq 6$ | 3.76 | 0.65 | 0.72 |

Data is weekly and in logged levels (from 2 January 1999 to 3 July 2009). Source: Bloomberg
† The Reimers (1992) finite sample correction has been applied to all statistics

The results here are starkly different from the pairwise approach: I find there to be no cointegrating vectors at the 5% level for either set of indices after I apply the finite sample correction (68), which in this case is $\frac{545-(3)(7)}{545} \simeq 0.9617$. While these results maybe somewhat difficult to rationalize with the bivariate results, it is important to bear in mind that the Johansen (1988) procedure does not allow for a structural break. Moreover, Gregory *et al.* (2004) found that the correlation of $p$-values between a pairwise "residual-based test and a system-based test is very low even as the sample size gets large." Also, recall that the results from the pairwise analysis were *not* indicative of highly related markets (*i.e.,* the greatest number of bivariate cointegrating vectors I find from the pairwise analysis is 5 out of a possible 21). There is also evidence of structural breaks in the data since the Zivot and Andrews (1992) procedure finds more cointegrating vectors than the Engle and Granger (1987) procedure for both types of indices. Nonetheless, the results of the systems approach indicate that any market in the G7 would offer investors gains in terms of diversification.

## 10 Conclusion

This essay focuses on the possibility that investor home bias is actually a rational choice rather than a pathology. To examine this possibility, I develop a bivariate model of two unit root stock prices and find that when the two prices are cointegrated investors optimally shift their holdings towards one asset. This polarization simultaneously reduces their level of diversification. In other words, in the case of a common stochastic trend, the cost of a home bias to investors (particularly long term investors) is less than indicated by previous studies. Importantly, the results are robust for both the traditional portfolio problem and the exponential CARA utility function.

These results, however, do not explain home bias. Indeed, the puzzle is not just in explaining why people in a given country are not very well diversified internationally. One must also explain why investors in other countries are also not diversified internationally in a completely different way (*i.e.,* one must explain why all investors hold mostly domestic assets). For example, from Canada's perspective home bias would imply that Canadian investors hold mostly Canadian assets. Such holdings could easily be explained if either (i) expected returns on Canadian assets were greater than on those of foreign assets or (ii) Canadian assets were lower risk than foreign assets. In this example, the home bias puzzle is then why foreigners do not similarly own mostly Canadian assets. Researcher, like French and Poterba (1991), have argued that the perceived risk and return on domestic and foreign assets may vary from nation to nation (like the cases explored in Section 5.2) and thus explain why each country's investors seem to be home biased. I expand on this hypothesis and show that cointegrated markets actually magnify the impact of varying perceptions of risk and return and make home bias easier to explain. In particular, I show that perhaps French and Poterba (1991) overstated the high implied expected returns for home markets (*i.e.,* investor expectations about home markets were not as wrong as French and Poterba (1991) concluded).

Thus, with the above theoretical results in mind, I conduct tests to determine if the G7 markets were actually cointegrated for the period of 1999 to 2009. Specifically, I examine empirically the possibility that home bias is less costly (and/or puzzling) than originally reported by French and Poterba (1991) and others. My findings, however, are not indicative of highly integrated markets. In particular, I found few cointegrating vectors, implying that a home bias may still be quite expensive to investors, and thus still puzzling to researchers. Specifically, under the systems method, I found zero cointegrating vectors, which would indicate that all of the series are driven by separate stochastic trends. Thus, while home bias can sometimes be an optimal response for investors, in the case of the G7 markets it appears that for the last ten years it has been more of a pathology.

A logical extension of my research would be to obtain portfolio data and use it directly to test the theory I develop above. However, such data is extremely difficult to obtain. Similarly, analyzing the trends in home bias over the past decade would also be insightful given my findings in this essay.

## References


Allen, D. E. and G. MacDonald (1995, February). The long-run gains from international equity diversification: Australian evidence from cointegration tests. *Applied Financial Economics 5*(1), 33–42.

Baele, L., C. Pungulescu, and J. Ter Horst (2007). Model uncertainty, financial market integration and the home bias puzzle. *Journal of International Money and Finance 26*(4), 606–630.

Chang, T. (2001). Are there any long-run benefits from international equity diversification for Taiwan investors diversifying in the equity markets of its major trading partners, Hong Kong, Japan, South Korea, Thailand and the USA. *Applied Economics Letters 8*(7), 441–446.

Cheung, Y. W. and K. S. Lai (1993). Finite-sample sizes of johansen's likelihood ratio tests for cointegration. *Oxford Bulletin of Economics & Statistics 55*(3), 313–329.

Cochrane, J. (1991). A critique of the application of unit root tests. *Journal of Economic Dynamics and Control 15*(2), 275–284.

Coeurdacier, N. (2009). Do trade costs in goods market lead to home bias in equities? *Journal of International Economics 77*(1), 86–100.

Danthine, J. and J. Donaldson (2005). *Intermediate financial theory.* Burlington, MA: Academic Press.

DeFusco, R. A., J. M. Geppert, and G. P. Tsetsekos (1996, May). Long-run diversification potential in emerging stock markets. *The Financial Review 31*(2), 343–63.

Dickey, D. and W. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association 74*(366), 427–431.

Enders, W. (1995). *Applied econometric time series* (1 ed.). New York, NY: Wiley.

Engle, R. and C. Granger (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica*, 251–276.

Fama, E. (1965). The behavior of stock-market prices. *Journal of business 38*(1), 34.

French, K. R. and J. M. Poterba (1990, December). Japanese and u.s. cross-border common stock investments. *Journal of the Japanese and International Economies 4*(4), 476–493.

French, K. R. and J. M. Poterba (1991, May). Investor diversification and international equity markets. *American Economic Review 81*(2), 222–26.

Granger, C. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics 16*(1), 121–130.

Gregory, A., A. Haug, and N. Lomuto (2004). Mixed signals among tests for cointegration. *Journal of Applied Econometrics 19*(1), 89–98.

Grubel, H. (1968). Internationally diversified portfolios: welfare gains and capital flows. *The American Economic Review*, 1299–1314.

Grubel, H. and K. Fadner (1971). The interdependence of international equity markets. *Journal of Finance*, 89–94.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control 12*(2-3), 231 – 254.

Juselius, K. and D. F. Hendry (2000, December). Explaining cointegration analysis: Part ii. Discussion Papers 00-20, University of Copenhagen. Department of Economics.

Kanas, A. (1998). Long-run benefits from international equity diversification: a note on the Canadian evidence. *Applied Economics Letters 5*(10), 659–663.

Kanas, A. (1999). A note on the long-run benefits from international equity diversification for a UK investor diversifying in the US equity market. *Applied Economics Letters 6*(1), 49–53.

Karlsson, A. and L. Nordén (2007). Home sweet home: Home bias and international diversification among individual investors. *Journal of Banking and Finance 31*(2), 317–333.

Kasa, K. (1992). Common stochastic trends in international stock markets. *Journal of Monetary Economics 29*(1), 95–124.

Kho, B., R. Stulz, F. Warnock, and K. Sinlim-Dong (2006). Financial globalization, governance, and the evolution of the home bias. *NBER Working Paper*.

Levy, H. and M. Sarnat (1970). International diversification of investment portfolios. *The American Economic Review*, 668–675.

Lewis, K. K. (1996). What can explain the apparent lack of international consumption risk sharing? *The Journal of Political Economy 104*(2), 267–297.

Lewis, K. K. (1999). Trying to explain home bias in equities and consumption. *Journal of Economic Literature*, 571–608.

Lo, A. and A. MacKinlay (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of financial studies*, 41–66.

Maddala, G. and I. Kim (1998). *Unit roots, cointegration, and structural change*. Cambridge Univ Pr.

Markovitz, H. (1959). *Portfolio selection: Efficient diversification of investments*. New York: John Wiley.

MSCI (2009). Index definitions. `http://www.mscibarra.com/products/indices/equity/definitions.jsp`.

Narayan, P. and R. Smyth (2004). Modelling the linkages between the Australian and G7 stock markets: common stochastic trends and regime shifts. *Applied Financial Economics 14*(14), 991–1004.

Obstfeld, M. and K. Rogoff (2000). The six major puzzles in international macroeconomics: Is there a common cause? *NBER Macroeconomics Annual 15*, 339–390.

Panton, D., V. Lessig, and O. Joy (1976). Comovement of international equity markets: A taxonomic approach. *Journal of Financial and Quantitative Analysis*, 415–432.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica*, 1361–1401.

Phillips, P. and P. Perron (1988). Testing for a unit root in time series regression. *Biometrika 75*(2), 335–346.

Poterba, J. and L. Summers (1989). Mean reversion in stock prices: Evidence and implications. *NBER working paper*.

Reimers, H. E. (1992, December). Comparisons of tests for multivariate cointegration. *Statistical Papers 33*(1), 335 – 359.

Ripley, D. (1973). Systematic elements in the linkage of national stock market indices. *Review of Economics and Statistics*, 356–361.

Roll, R. (1992). Industrial structure and the comparative behavior of international stock market indices. *Journal of finance*, 3–41.

Samuelson, P. (1969). Lifetime portfolio selection by dynamic stochastic programming. *The Review of Economics and Statistics*, 239–246.

Sims, C. (1988). Bayesian skepticism on unit root econometrics. *Journal of Economic Dynamics and Control 12*(2/3), 463–74.

Sørensen, B., Y. Wu, O. Yosha, and Y. Zhu (2007). Home bias and international risk sharing: Twin puzzles separated at birth. *Journal of International Money and Finance 26*(4), 587–605.

Taylor, M. and I. Tonks (1989). The internationalisation of stock markets and the abolition of UK exchange control. *The Review of Economics and Statistics*, 332–336.

Tobin, J. (1958). Liquidity preference as behavior towards risk. *The Review of Economic Studies*, 65–86.

Zivot, E. and D. W. K. Andrews (1992, July). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics 10*(3), 251–70.