NBER WORKING PAPER SERIES

INTERNATIONAL RISK SHARING IN THE SHORT RUN AND IN THE LONG RUN

Marianne Baxter

Working Paper 16789
http://www.nber.org/papers/w16789

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2011

I am grateful for stellar research assistance from Hirozaku Ishise and Timothy Layton. Helpful comments were provided by Akito Matsumoto, participants in the CEPR Conference on International Risk-Sharing, seminar participants at the Federal Reserve Bank of Boston, the Federal Reserve Bank of New York and the members of the NBER IFM group. Remaining errors are my own. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

International Risk Sharing in the Short Run and in the Long Run
Marianne Baxter
NBER Working Paper No. 16789
February 2011
JEL No. E2,E21,E32,F11,F15,F2,F4,F41

ABSTRACT

International risk-sharing has far-reaching implications both for economic policy and for basic research in economics. When countries do not share risk, individuals in those countries experience fluctuations in their consumption levels that are undesirable and possibly unnecessary. This paper extends and refines the study of international risk-sharing in two dimensions. First, this paper investigates risk-sharing at short vs. long horizons. Countries might, for example, pool risks associated with high-frequency shocks (e.g., seasonal fluctuations in crop yields) but might not share risks associated with low frequency shocks (e.g., different long-run national growth rates). Second, this paper studies bilateral risk-sharing, which is different from the approach taken in most previous studies. We find that there is evidence of substantial international risk-sharing at medium and low frequencies. There is evidence of high and increasing risk-sharing within Europe that is not apparent for other regions of the world.

Marianne Baxter
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and NBER
mbaxter@bu.edu

International capital markets are becoming increasingly integrated across countries. Their development throughout the last half of the 20th century was aided by rapid advances in communications and computing technology. It is natural, therefore, to ask whether the increase in "globalization" has led to an increase in international risk-sharing. This paper contributes to the large existing literature on international risk-sharing by investigating two facets of risk-sharing that merit closer attention. The first is the temporal dimension to risk-sharing. Countries might, for example, pool risks associated with high-frequency shocks (e.g., seasonal fluctuations in crop yields or demands for raw materials) but might not share risks associated with low frequency shocks (e.g., different long-run national growth rates). This paper explicitly investigates risk-sharing at different horizons. Second, this paper studies bilateral risk-sharing and the way in which bilateral risk-sharing varies with the risk-sharing horizon. Due to factors such as financial linkages, common cultural linkages, or simply proximity, countries might share risks with some countries, but not with others. By studying bilateral risk-sharing, we may be able to determine which economic or other forces are associated with greater risk-pooling.

This paper builds on and is inspired by many contributions in this research area. Early tests were conducted by Obstfeld (1994), Lewis (1996,1999) and Canova and Ravn (1996). Building on that early work, the literature has developed several interesting strands of research, including a comparison of within-country vs. cross-country risk sharing; the effects of European unification on risk-sharing; the effect of country size and openness on risk-sharing, and a host of other related questions.[1]

1 A Framework for Studying International Risk-Sharing

This section develops a framework for studying international risk-sharing. Most of the existing research on risk-sharing has been based on variations of this framework. Other variables may also enter the utility function.

1.1 Time-separable, constant-relative-risk-aversion preferences

The most common specification of utility employed in the risk-sharing literature is the specification in which each agent wishes to maximize a time-separable, CRRA utility as a function of aggregate consumption. The discount factor and the rate of relative risk aversion are assumed to be identical across countries.

[1] A review of the recent literature can be found in Kose, Prasad and Torrones (2007). See also Becker and Hoffman (2006) and Flood, Marion and Matsumoto (2009) for recent papers that discuss aspects of long-run risk sharing.

The country-i expected utility function is:

$$U_i = E_0 \sum_{t=0}^{\infty} \beta^t u\left(c_{it}\right) = E_0 \sum_{t=0}^{\infty} \beta^t \frac{c_{it}^{1-\sigma}}{1-\sigma}.$$

This problem of finding an optimal allocation of consumption across countries can be found by solving a planner's problem. The planner assigns a weight, θ_i, to utility in country $i = 1, 2, ..., I$. The world social welfare function at time t is thus:

$$W_t = \sum_{i=1}^{I} \theta_i U_{it}.$$

Each country is populated by a large number of identical, infinitely-lived individuals, η_i .The world resource constraint is:

$$\sum_{i=1}^{I} \eta_i c_{it} \leq \sum_{i=1}^{I} \eta_i y_{it} = f(\varsigma_t)$$

where ζ_t denotes exogenous shocks at date t. We will not focus on the process by which consumption goods arrive on the planet. They could arrive as stochastic endowments in each country; under this interpretation ζ_t would be a vector of country-specific endowments. Alternatively, output could be produced using capital and work effort. Under this interpretation, ζ_t would be a vector of productivity shocks. For the present purpose, it is enough that output has a stochastic component and can be allocated contemporaneously by the planner.

Let λ_t be the current-valued multiplier for the world resource constraint. The first-order condition for the planner's optimal assignment of consumption goods to country i at time t is:

$$\eta_i \lambda_t = \theta_i \frac{\partial u(c_{it})}{\partial c_{it}} \tag{1}$$

This equation has formed the basis for many empirical investigations of aspects of risk-sharing. In the next section, we describe two main approaches to testing risk-sharing and discuss the advantages and disadvantages of each.

2 Methodology and Prior Research

This section addresses the methodological issues involved in testing risk-sharing or, indeed, testing any other prediction of an economic model as summarized in an Euler equation or functions of Euler equations. We will distinguish between 'direct' tests and 'indirect' tests, as elaborated below. Subsequently, we address the relationship between tests of risk-sharing and tests of the permanent income hypothesis.

2.1 A Direct Test of Risk-Sharing

A direct approach to testing risk-sharing is to test the risk-sharing condition implied by the first-order conditions. The ratio of the first order conditions at date t for countries i and j (one of which could be the "world") yields the following implication of complete risk-sharing:

$$\frac{\eta_i}{\eta_j} = \frac{\theta_i c_{it}^{-\sigma}}{\theta_j c_{jt}^{-\sigma}}$$

This equation is not ready to take to the data for two reasons. First, the planner weights θ_i are not known. Second, the time-series for aggregate consumption shows strong evidence of I(1) behavior leading to problems of estimation and inference if consumption is used in levels. Therefore, the risk-sharing condition is usually expressed in a log transformation that emphasizes changes in consumption, not levels of consumption:

$$(\log c_{i,t+k} - \log c_{it}) = (\log c_{j,t+k} - \log c_{jt}). \tag{2}$$

For the class of CRRA preferences which depend on consumption alone, direct testing of risk-sharing includes testing whether the k-period differences of log consumption are perfectly correlated across countries. Of course, no economic relationship ever holds in the data without error. The deviation from the implied function of the first-order conditions must therefore be given some economic interpretation. A natural interpretation might be that it represents measurement error. As such, the error should have mean zero (which it will, by construction, in a regression setting); it should also be serially uncorrelated absent issues of time-aggregation that could induce persistence in the error. A regression estimate of (2) should lead to a high R^2. Alternatively, the correlation between consumption growth in countries i and j should be very high; it will be different from 1 (under this interpretation of the regression residual) only to the extent that there is measurement error.

2.2 An Indirect Test of Risk-Sharing

An indirect test of the empirical validity of the Euler equation tests whether variables that ought not to be correlated with the error in the Euler equation do, in fact, explain a significant amount of variation in the error. Although a bilateral test could be formulated, more typically the researcher begins with a transformation of equation (1), as follows:

$$\ln \eta_i + \ln \lambda_t = \ln \theta_i + \ln \left(\frac{\partial u(c_{it})}{\partial c_{it}} \right). \tag{3}$$

The world marginal utility of consumption, λ_t, is proportional to the world consumption level, c_t^w, which means that:

$$\ln c_{it} - \ln c_t^w = \ln \theta_i - \ln \eta_i. \tag{4}$$

Thus the country i consumption level differs from the world consumption level only to the extent that the planner weight is different from the country's share of world population. In any case, fluctuations in country i consumption will move in lockstep with world consumption. The model implies that $\ln c_{it} - \ln c_t^w$ is stationary so there is no reason to filter the data to achieve stationarity.

Early examples of indirect tests in the context of risk-sharing can be found in the work of Obstfeld (1994) and Lewis (1996), which in turn are based on earlier work by Cochrane (1991) and Mace (1991). Each of these contributions began with the observation that consumption growth in a particular country should be uncorrelated with country-specific risks if risk-sharing is complete. In implementing this test, Lewis used the deviation of lagged national output from global output as a measure of country-specific risk that should carry a regression coefficient of zero. Thus the simplest version of her tests involves running the following regression:

$$\ln c_{it} - \ln c_t^w = \alpha + \beta(\ln y_{i,t-1} - \ln y_{t-1}^w) + \epsilon_{it}$$

Lewis tested risk-sharing by testing the null hypothesis $\beta = 0$.

Obstfeld's 'workhorse' equation is the first-differenced version, below, because he was concerned about stochastic trends (he did not want to embed the cointegration assumption of the null hypothesis):

$$\Delta c_{it} = \alpha + \gamma \Delta c_t^w + \epsilon_{it}.$$

Obstfeld estimated this equation on Penn World Tables data for 1950-1988. Part of his paper is devoted to calculation of the correlation between Δc_{it} and Δc_t^w. He finds higher correlations in the 1973-1988 period than in the earlier, 1950-1972 period. This would be interpreted as a "direct" test of risk-sharing.

Obstfeld then explores whether other variables, such as GDP growth and government spending growth explain idiosyncratic variation in c_{it}. His investigation into these other variables is partly motivated by a desire to see whether these other variables can explain (in an R-squared sense) consumption growth, and whether they are plausibly stand-ins for preference shocks. He also argues, convincingly, that government spending shocks are uninsurable internationally given the concomitant moral hazard issues, and provides some support for this hypothesis. These tests are "indirect" because they test whether something that should not explain consumption movements does, in fact, have predictive content for consumption. If there is complete risk-sharing, the country-specific variables, such as the national component of output, should not matter for national consumption.[2]

An important problem with the indirect approach to testing risk-sharing is that every single economic variable that does not enter the Euler equation should be uncorrelated with the growth rate of marginal utility. The candidates include state and local fiscal variables; monetary policies; exchange rates, unemployment... the list is endless. In fact, foreign "country-specific" variables should not matter either. It is an impossible task to collect the data and then test whether every one of these variables carries a zero coefficient. One can never, with this approach, claim to have verified the existence of risk-sharing. All that can be said is that, with respect to a particular set of variables, there is (or is not) strong evidence against risk-sharing.

2.3 Risk-Sharing and the Permanent Income Hypothesis

This section considers the relationship between tests of international risk-sharing and tests of the permanent income hypothesis. Many tests of risk-sharing, both across countries and across individuals, have been conducted using specifications very similar to tests of permanent income theory. A typical test of the permanent income theory regresses consumption growth on (i) a measure of the real rate of interest, capturing the effect of intertemporal substitution on consumption growth, and (ii) the lagged growth rate of income. The null hypothesis implied by permanent income theory is that lagged income growth should carry a zero coefficient. Typically, this null hypothesis is rejected by the data. Cochrane (1991) and Mace (1991) used similar specifications to study interpersonal risk-sharing. Lewis (1996) uses a modification of this approach to study international risk-sharing. In each case, the test of risk-sharing amounted to a test that the coefficient on lagged income growth was zero.

This raises a natural question: 'Is the concept of risk-sharing inextricably linked with the concept of the permanent income hypothesis?' The answer is 'no,' as a simple example will illustrate. Suppose that there are two countries, and output is produced using capital and labor. Each period, a constant fraction of world output is allocated to investment, and the rest is allocated to consumption. The investment and consumption amounts are allocated to the two countries according to a constant sharing rule; say, $\frac{1}{2}$ for each country. In this situation, the risk-sharing condition (2) would hold. However, so long as there is persistence and cross-country correlation in output growth (which is true in the data), the test of the PIH would fail, since lagged output growth would predict current-period output growth which would, in turn, be correlated with current-period consumption growth. Thus risk-sharing does not imply the

[2] Although most versions of Lewis's tests rejected risk-sharing, she found that she could not reject risk sharing when conditioning on no capital market restrictions and including the effects of nonseparabilities.

PIH. Showing that the PIH does not imply risk-sharing is straightforward; the two countries can simply behave as autarky economies, but following the PIH within each country. These examples show that there is no necessary connection between the permanent income hypothesis and risk-sharing. Either of these hypotheses can be true in the absence of the other. This is fortunate, since many economists believe that there is abundant empirical evidence against the PIH. If risk-sharing required that the PIH hold, there would be little point to studying risk-sharing.

3 Empirical Results

This sub-section presents results for direct tests of risk-sharing in the case in which preferences are of the constant-relative-risk-aversion, time-separable form. This is the specification that has received the most study in prior investigations. Our data are quarterly, from 1960:1 through 2009:2 for 21 countries. Nominal variables are translated to constant US dollars using the nominal exchange rate vis-a-vis the US together with the US GDP deflator. The risk-sharing condition says that k-period growth rates of consumption should be equalized across countries for all values of k. That is: these k-period growth rates should be perfectly correlated across all pairs of countries. The majority of empirical risk-sharing papers focus on the k=1 case, i.e., the data are first-differenced. The reason for using the first-difference filter is rarely explicit (Obstfeld (1994) is an exception). One likely reason is that first-differencing achieves stationarity of an I(1) consumption series. However, if consumption is I(1), differencing at any horizon will achieve stationarity, as will any of a large set of common filters such as one- or two-sided moving averages, the Hodrick-Prescott filter, or the band-pass filters of Baxter and King (1999). Some of these filters have desirable properties while others, such as the Hodrick-Prescott filter with asymmetric treatment of the endpoints, are well known to distort the data. In the present paper, we make contact with current practice by studying growth rates of marginal utility. However, we extend the typical focus of quarterly differences to study differences extending to a maximum of 24 quarters.

We begin by illustrating the typical pattern of correlation between consumptions across countries and the relationship between bilateral consumption correlations and output correlations. The first fact that emerges from the data is that consumption correlations for first-differenced data (k=1) are typically the lowest of all correlations computed for various values of k. The difference can be substantial. Figure 1 shows a scatter plot of consumption correlations for all pairs of countries (210 pairs) for the k=1 case and the k=16 case (k=16 corresponds to four-year differences–16 quarters). This figure shows that there is just one country pair for which the quarterly first-difference correlation is just a tiny bit larger than the four-year difference correlation. For the 209 other pairs of countries, the correlation of 16-quarter differences is substantially larger than the correlation of first differences. The median first-difference correlation is 0.50; the median for the 16-quarter differences is 0.78.

3.1 Trading partner correlations

Figure 1 also indicates the bilateral correlation for each country with its largest trading partner as a solid, filled "o" on the graph. One might expect that consumption correlations with one's trading partner would plausibly be higher than consumption correlations with other countries because international trade in goods is one important way that countries can smooth marginal utility. The figure suggests that trading partner correlations are not obviously different from the distribution of all pairwise correlations.

We wish to formally test the null hypothesis that the correlation of a given country with its major trading partner is not drawn from a different distribution from the correlations with all other countries in the sample (holding k fixed). Because of the small sample size and because the distribution requirements for use of a t-statistic are not satisfied in this context, we used the Mann-Whitney-Wilcoxon U-test (see Mann-Whitney (1947), Wilcoxen (1945)). In our context, where we are comparing the correlation of one partner country ($n_1 = 1$) with the correlations of the other nineteen countries in the sample ($n_2 = 19$) the U-statistic on which this test is based is simply the number of countries in sample n_2 with correlations greater than the correlation with the partner country. The z-statistic below is asymptotically standard normal for 'large' samples (greater than about 20):

$$z \equiv \frac{U - \mu_u}{\sigma_u}$$

where

$$\mu_u = n_1 n_2 / 2$$

and

$$\sigma_u = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}.$$

Table 1 presents the results of this test. The table lists each country in the sample in the first column, followed by its largest trading partner in the second column. The remaining columns indicate the results of the U-test for differencing horizons k=1, 4, 8, 12, 16, 20, and 24.[3] An asterisk, *, indicates that the trading partner correlation is significantly larger than the other pairwise

[3] Note that the test is not symmetric, i.e., country A can be more correlated with its largest trading partner, country B, than with A's other trading partners while country B may not be more highly correlated with A, its largest trading partner, than it is with B's other trading partners.

correlations at the 10% significance level.[4] The top part of the table presents results for the large group of countries whose largest trading partner is Germany. Austria, Luxembourg, the Netherlands, and to a smaller extent, Belgium, show significantly higher correlations with Germany than with the rest of the countries in the sample. These are four of the smallest countries in the group of German trading partners and also are among the closest to Germany geographically. Germany's larger or more distant trading partners–Finland, France, UK, Italy, and Sweden–do not show significantly higher correlations with Germany than with other countries.

Among the rest of the European countries (those for which Germany is not the largest trading partner), we find some countries that have significantly higher correlations with their main trading partner. These are Spain (with France); Ireland (with the UK); and Portugal (with Spain). We do not find higher trading partner correlations for Germany (with France); Norway (with the UK), and Iceland (with the Netherlands). These findings reinforce the results above, where Germany was the main trading partner, that smaller countries tend to have higher correlations with their largest trading partner if that trading partner is also nearby.

When we look at the Australiasia group, the same finding emerges. New Zealand is more highly correlated with Australia than with other countries in our sample, but Australia and Japan are not more highly correlated with their respective trading partners than with the group of countries as a whole. New Zealand is small relative to Australia and is also much closer to Australia than any other trading partner. For North America, however, the pattern is different. The two large countries, US and Canada, are highly correlated with each other (the US is more highly correlated with Canada than with other trading partners at all values of k), but Mexico and Canada are more correlated with the US only for low values of k. This may be consistent with the evidence above if we consider that the US and Canada, while both large countries, have no other large trading partners nearby as do France, Germany, Italy, etc. in Europe. This could explain why the US-Canada correlations are high, but the correlations between large European countries are not particularly high relative to other trading partners. Also, Mexico is an outlier in this sample of countries, having consumption and output movements much less correlated with the world cycle than the other countries. Thus the finding that Mexico's comovement with the US is higher only at low frequencies is understandable.

[4] Because of the structure of our hypothesis (with $n_1 = 1$), the asterisk also indicates that the trading partner correlation exceeded all but 2 or fewer of the other correlations. It should be carefully noted that this test does not take into account the serial correlation in the k-differences of consumption that are used to construct the cross-country correlation coefficients.

3.2 Kernel estimates of the distribution of correlations

Another way to gauge the effect of the differencing interval, k, on consumption correlations, is to look at the distribution of correlations as k varies. Figure 2 plots a kernel density of the pairwise consumption correlations for k=1,2, 16, 24. The kernel density for k=1 has substantial mass at low correlations, roughly between zero and 0.5, and also has a mode at about 0.8. The country pairs represented in the region of the right-hand mode (correlations about 0.8) are mainly the European countries. As the differencing interval rises, the distributions of consumption correlations show increasing mass at high correlations and much less mass at low correlations. For k=1, the modal correlation is 0.50. By contrast, for the longest differencing interval that we consider, k=24 quarters, the modal correlation is 0.85.

Why are consumption correlations so low for the k=1 case compared with longer horizons? The first reason that springs to mind is that this finding is an artifact of the first-difference filter. The first-difference filter downweights low frequencies and business-cycle frequencies while putting greater than one-for-one weight on high frequencies. Specifically, the reweighting of the first-difference filter toward low frequencies–the 'noise' frequencies–will result in a low measured correlation for first-differenced data. This was illustrated by Baxter and King (1999). A graphical illustration of this effect is shown in Figure 3.

The use of the first-difference filter on highly persistent time series will lead to low correlations of the filtered data. This was demonstrated by Baxter (1994) in the context of the relationship between real exchange rates and real interest rates, and was shown for a large number of macroeconomic time series by Baxter and King (1999). The fact that we find low correlations in the present context is to be expected. This finding is not necessarily a sign of low risk-sharing. It is more likely an artifact of reweighting the time series to emphasize noise components.

There is also the possibility of measurement error in the data. The presence of classical (white noise) measurement error in the consumption series would naturally lead to a low correlation of first-differenced data. The next sub-section explores this possibility.

3.3 Measurement error: An example

Classical measurement error is white noise added to the 'signal'–the true measure of consumption. To take a concrete example, suppose that consumption in each country follows identical random walk processes, so that true consumption growth is perfectly correlated across countries at all horizons. To this random walk process, add independent white noise error to each country-level observation. The innovation standard deviation of the white noise (the "measurement error") is set equal to .25 times the innovation standard devation of the consumption process.

Next, simulate this economy for 200 periods–approximately the same length as the available sample length of 198 periods. Then compute the correlation coefficients of k-differences of the simulated processes, for k=1 to 40. Although the true correlation is 1 for all values of k, the effect of this measurement error is to reduce the value of the correlation coefficient for small values of k, as shown in the top panel of Figure 4.

As Figure 4 shows, the effect of the measurement error is most pronounced for k=1. The influence of these white noise processes on the correlation of the k-differenced data dies out by about k=8, or about 2 years for quarterly data. This panel also shows a 95% confidence interval constructed using den Haan and Levin (1997, 2000) VAR-HAC estimates of the standard deviation of the correlation coefficient at each value of k. The measurement error has its largest effect at the $k = 1$ horizon, as expected, since the measurement error is white noise and is amplified by the first-difference filter as demonstrated above. What is more surprising is that the white-noise measurement error continues to affect the estimated correlations at longer horizons. Although the true correlation is 1.0, the correlation of the series with measurement error remains below 0.98 until $k > 10$. That is: the measurement error affects measured correlations at horizons of two to three years which is within typical definitions of business-cycle frequencies. Thus, measurement error influences high frequencies most profoundly, but also affects lower frequencies in a way that is somewhat surprising.

The confidence intervals implied by the 2-standard-error bands in the top panel of Figure 4are noticeably jagged. This is an artifact of the implementation of the VAR-HAC procedure in a small sample characterized by high persistence in the underlying process.

Conceptually, the idea is the same as finding the variance of the following process:

$$x_t = \psi x_{t-1} + u_t$$

with $0 < \psi < 1$ and $u_t \tilde{} i.i.d.\ N(0, \sigma^2)$. The variable x is zero mean with variance $\sigma^2/1 - \psi^2$. The estimate of the variance of x will be sensitive to small changes in the value of ψ, especially if the process is highly persistent. In the context of our k-differenced consumption data, we have two factors leading to large standard errors and also volatile estimates of the standard errors as k changes. First, the process for consumption is itself highly persistent. Second, even if consumption itself were white noise, the k-differences would be highly persistent because each of the observations for a particular differencing horizon, k, contains (k-1) of the same sample points as the preceding observation. Thus, persistence is built in via the differencing process. In a small sample, the estimated persistence of a component of the correlation coefficient, say the mean, can change dramatically as k changes. This will translate into highly volatile standard errors. We will observe this volatility in 'real-world' data for the same reasons that we observe it in the artificial data presented above.

The bottom panel of Figure 4 shows the estimated correlations and confidence interval for a sample of 20,000 observations. The jaggedness in the confidence intervals disappears. In fact, the standard deviations have shrunk to nearly zero as we would expect for such a large sample. However, the low estimated values of the correlation coefficient for small values of k does not disappear since it depends on the innovation variance of the measurement error, not the sample size.

3.4 Confidence intervals for bilateral correlations

This sub-section presents results for consumption correlations for each country against two natural benchmarks: (i) world consumption; and (ii) consumption of the country's largest trading partner. All variables continue to be in per capita terms in real US dollars. If there is complete risk-sharing, individual country growth rates of consumption will be correlated with the growth rate of world consumption at all horizons. If there is a lack of complete risk sharing, or if risk sharing is complete only at certain frequencies, then we may observe that the consumption correlations differ across horizons. In a world with costly trade, it may be the case that risk sharing is higher with a country's major trade partner than it is with the world as a whole. If so, this finding could also reflect industrial structure or linkage in financial institutions, as investigated empirically in Baxter and Kouparitsas (2003,2005). However, it remains an interesting empirical question whether a country's correlation is higher with the world as a whole or its one's largest trading partner.

As shown above, the confidence intervals for the correlations are likely to be wide and quite jagged. A major problem arises when there is a high correlation between two series since the sample estimate of the correlation coefficient does not have a normal distribution, being bounded in the interval [-1,1]. Under the typical assumption on the normality of the error terms in the driving process for consumption, the standard error bands for consumption using the VAR-HAC standard errors can exceed the theoretical bounds of [-1,1]. For this reason, we construct the Fisher (1915, 1921) z-transform of the correlation coefficient. This statistic was originally designed precisely to produce a normal distribution for a monotonic transformation of the correlation coefficient.[5] The Fisher $z-$transformation is as follows, where r denotes the sample estimate of the correlation coefficient:

$$z = \frac{1}{2}\ln\frac{1+r}{1-r} = \operatorname{arctan} h(r)$$

[5] Akito Matsumoto suggested the use of the Fisher z-transform.

Fisher (1915, 1921) showed that, if variables X and Y are have a bivariate normal distribution, then z is approximately normally distributed and is unbiased for the population correlation coefficient, ρ. We therefore apply the VAR-HAC approach to estimation of the mean and standard error of z, constructing two-standard-error bands for this variable, and then reporting the implied mean and standard error bands for ρ using the inverse of the function above that defines z:[6]

$$r = \frac{\exp(2z) + 1}{\exp(2z) - 1}.$$

Figure 5 presents results on risk-sharing for several of the countries in our sample. The countries in the sample are taken from three broad geographic areas: Europe; North America; and Australasia. Considering Europe first, we find that the correlations of the consumption growth rates of European countries with their largest trading partner (most frequently Germany or France) tends to be higher than their correlation with the world aggregate. This also tends to be true at all horizons. However, the standard error bands (shown only for the world aggregate) are large enough that one could not say that there is a significantly higher correlation with the trading partner than with the world. The results are nevertheless suggestive of a stronger degree of risk-sharing within Europe than for European countries vis-a-vis the world as a whole. We take a closer look at this in the next sub-section.

We do not find a similar pattern of strong geographic risk-sharing for the other two groups. The correlations among the three North American countries are much lower than for the European countries, even for Canada vs. the US. The correlation between Canadian consumption growth and that of the US is only about 0.50 for the first-differenced data, but rises to about 0.80 for long differences. Mexico's correlation with the US is very low–only about 0.25 for all horizons. There is no evidence of strong North American risk-sharing as was found for Europe.

The pattern of low regional risk-sharing continues with the Australia/New Zealand/Japan group–admittedly an odd combination when one considers the large distances between the first two countries and Japan. Still, Japan is Australia's largest trading partner in our sample (China is currently Australia's largest trading partner, but China is not in our sample). One might therefore expect that Australia's consumption correlations would be more correlated with Japan's than with the world's. This is not the case, however. For the first-differenced data, the correlation is only about 0.20, rising to about 0.60 for the longest horizons considered. Australia's correlation with the world cycle is substantially higher–about 0.55 at k=1, rising to nearly 0.90 for k=24. New Zealand, by contrast, shows consumption correlations more highly correlated with its largest trading partner, Australia, than it does with the world cycle.

[6] The confidence interval for z is symmetric but the implied confidence interval for ρ is not.

Japan's largest trading partner is the US (not part of Asia, so the regional link is broken here.) We find that Japan's correlation with the world is higher than it is with the US–substantially higher for short horizons. For the k=1 case, Japan's correlation with the US is only 0.20, while its correlation with the world cycle is nearly 0.70. The gap narrows as k rises, but at the longest horizon considered, the correlation with the world is just over 0.90, while the correlation with the US is about 0.75.

3.5 Risk-sharing within Europe

We can formally test the hypothesis that risk-sharing between pairs of European countries is higher than between pairs of countries in which one or both lies outside of Europe. The U-test described earlier can be applied to the two groups. There are 13 European countries (we drop Iceland from the European group for these calculations), and 7 non-European countries. This leads to a total of 91 within-Europe correlations and 119 correlations in which one or both countries are not in Europe. The U-statistic computed for the hypothesis that the European correlations are no different from the non-European correlations is significantly different from zero at the 1% level for all values of k. In fact, the European correlation exceeds the non-European correlation in 88%-100% of cases when we perform the pairwise calculations required for the U-test (the range 88%-100% covers all values of k).

4 Is it the exchange rate?

Thus we are left with the finding that there is substantial risk-sharing between trading partners, and substantial regional risk-sharing within Europe, but little evidence of regional risk-sharing within other geographic groups. One potential explanation that springs immediately to mind is the stability of the exchange between European countries in contrast to the more-volatile exchange rates between the countries of the non-European groups. To explore whether this possibility could explain the foregoing results, consider the following thought experiment. Suppose that we assume that there is purchasing power parity, so that the price of goods is the same in all currencies. We could then construct a measure of real consumption in each country by deflating each country's nominal consumption expenditure by an appropriate deflator. This would remove the effect of exchange rate volatility on measured consumption completely. The truth probably lies somewhere between the consumption measured in exchange-rate-adjusted terms and consumption measured in own-currency constant units. Without detailed sectoral information, we can only bound the results that would be obtained under the assumptions of (1) perfect tradability without purchasing power parity, so that the appropriate construct is the one previously used; or

(2) the assumption that local currency prices are correct either because goods are nontraded or because purchasing power parity holds. This experiment thus allows us to bound the results for risk-sharing for these two polar cases. Further, and perhaps most importantly, this analysis allows us to determine whether the European-risk-sharing finding is an artifact of exchange-rate stability in this region.

One interesting finding is that the exchange rate has an important effect on the distribution of correlations at varying horizons, as summarized earlier in the kernel densities. The kernel density estimates for real consumption, unadjusted for the exchange rate, is shown in Figure 6. One would probably have expected the effects of exchange rate volatility to be manifested at the high frequencies emphasized by low values of k since exchange rates are commonly viewed as approximately a random walk. However, what is really most striking about these kernel densities is that the exchange rate affects the longer differencing horizons, k=16 and k=24. Without the exchange rate adjustment, the distributions of the long differences show much more concentrated distributions at high correlations. The scatter plot of consumption correlations for k=1 and k=24 in this "PPP" case also illustrate this finding.The results for the "PPP" data continue to exhibit strong European risk-sharing, while most of the prior findings remain qualitatively similar. Figure 6 shows a set of graphs similar to those of Figure 4. We conclude that the strong intra-European risk-sharing is not an artifact of relatively smooth European exchange rates.

5 Is it recent?

Numerous empirical investigations of the character of national and international business cycles have documents changing patterns of statistical relationships over the past fifty years. For the purpose of this paper, we ask whether the two halves of our sample period exhibit different behavior with respect to estimated consumption correlations and their confidence intervals. The sample was split in half, with the first period encompassing 1960:1-1984:4, and the second period including 1985:2-2009:2. Shortening the sample to end in 2007:2 would not materially affect the results.

Rather than report tests for the full set of countries, we continue to focus on two places where changes in correlation would be important: (1) with the world aggregate; and (2) with a country's largest trading partner. For each subsample, the sample means of the k-differenced consumption data were calculated, and t-statistics were formed by subtracting the correlation from the second subsample from the first-period correlation, then dividing by the second-period variance-covariance matrix of the correlations. The first-period VCV could have been used instead with no important changes.

Table 1 presents the results for k=1, 4, 8, 16. There are several significant negative t-statistics for the k=1 case, but the majority of countries do not have significant differences with either the world or their major trading partner. There are only two positive t-statistics indicating higher correlation in the pre-1985 period. For k=4,8,16, significant differences across the time periods are rare. Thus, the finding of high risk-sharing is not a phenomenon driven by the most recent 25 years of data.

6 Is it risk sharing?

Certainly, if "risk sharing" means only highly correlated consumption growth rates, there is strong evidence for risk sharing at all but short horizons. The low correlations at short horizons may well be an artifact of measurement error, as shown earlier. However, if one means by risk sharing the deliberate pooling of consumption risks so that consumption profiles are more correlated across countries than would be the case in autarky, the case is much less clear. Consumption correlations are very tightly related to output correlations, as illustrated in Figure 7. (In this figure, consumption is measured in real US dollars.) The circles represent correlations with major trading partners. The first plot is for the k=1 case; the second is for k=16.

The point of these plots is to illustrate that the high correlation of consumption growth rates across countries might arise simply because outputs are so highly correlated. Myopic consumers who consume a fixed fraction of national GDP would look like individuals engaged in a high degree of risk sharing in a world where GDP growth rates are highly correlated. The hypothesis that consumption correlations are equal to output correlations for fixed k and fixed country pairs was tested using the HAC-VAR covariance matrix for consumption described earlier. The consumption covariance matrix was used because, a priori, it was expected to yield lower variance terms, thus biasing the test toward finding differences between consumption correlations and output correlations. In fact, the estimated covariance matrix was not very different for the output measure.

Of the 210 x 24 = 5040 test statistics constructed, fewer than 2% showed a significant difference between cross-country output growth rates and consumption growth rates. Spefically, only 17 country pairs yielded statistically significant (at the 1% level) t-statistics for the null of no difference between consumption and output correlations. Of these, nine had significant differences only for k<3. For all but two of the remaining country pairs, the significant differences disappeared between k=5 and k=10.

Why do the few significant correlations show up for low values of k? Marginal utility growth rates would likely be positively correlated—even with no risk-sharing at all—simply because of positive cross-country correlation in consumption and work effort stemming from the world component of the business cycle. Further, output contains high frequency fluctuations in investment that are not found in consumption, even though the measure of consumption contains purchases of durables. The smoother behavior of national consumption relative to national output at low frequencies, for the few countries displaying this pattern, could well be due to the combination of these two effects.

This sub-section began with a question: is it risk sharing? Specifically, is there evidence that consumption correlations are significantly higher than output correlations, reflecting the deliberate actions of individuals wishing to smooth consumption to a greater extent than is possible within their own borders? The evidence is quite negative. Although consumption correlation growth rates are strongly positively correlated for many country pairs at medium-to-long horizons, consumption correlations do not exceed output correlations. The conclusion from this analysis, then, is that there is little evidence of "marginal" risk-sharing, defined as consumption risk-sharing higher than the amount that would be achieved by autarkic economies smoothing consumption alone in the presence of capital accumulation technology.

7 Conclusion

This paper extends our understanding of international risk-sharing along both temporal and bilateral dimensions. We contrast the usual, indirect approach to risk-sharing to a direct approach based on computation of the correlation of growth rates of marginal utilities. The fact that most prior research uses the indirect approach probably stems from the longstanding use of this approach to studying the permanent income hypothesis. We show that the PIH and risk-sharing are independent theories and that neither implies the other.

Our results show that the post-1960 era has been characterized by very substantial risk-sharing between most countries and their major trading partners; between most countries and the world, and indeed between most countries and other individual countries. We find risk-sharing is lower at shorter horizons than longer horizons. Thus a focus on k=1 alone, as is the case in Obstfeld (1994), Lewis (1996) and the majority of subsequent analyses, would bias the results against finding support for risk-sharing. The strongest evidence in support of risk-sharing occurs at medium and low frequencies. The difficulty with looking at longer horizons is that, for many country pairs, the standard errors become extremely large. This difficulty will extend to any risk-sharing measure based on low-frequency movements in consumption.

To study the sensitivity of our findings to the emedded assumptions concerning tradability and the law of one price, we tested whether the results were substantially different when consumption was measured in real local currency units, rather than in real units of a single reference currency. The qualitative features of our results were not substantially affected. We also investigated whether the data were dominated by the recent 25-year period; this was not found to be the case. We then asked whether there was substantial evidence for "marginal" consumption risk-sharing: consumption movements more correlated than output movements. We found very little evidence to support this hypothesis.

Overall, our conclusion is that international risk sharing may be greater than you think. This paper's aim was to refocus attention on the extent of international risk sharing and to highlight the problems of inference when looking at measures of long run risk sharing.

References

[1] Baxter, Marianne (1994) "Real exchange rates and real interest differentials: Have we missed the business-cycle relationship?" Journal of Monetary Economics 33:1 (February 1994), 5-37.

[2] Baxter, M. and and R. King (1999) "Measuring business cycles: Approximate band-pass filters for economic time series", Review of Economics and Statistics 81:4, November 1999, 575-593.

[3] Baxter, M. and M.Kouparitsas (2003) "Trade Structure, Industrial Structure, and International Business Cycles," American Economic Review 51:2 (May 2003), 51-56.

[4] Baxter, M. and M.Kouparitsas (2005) "Determinants of Business Cycle Comovement: A Robust Analysis", Journal of Monetary Economics, 52:1 (January 2005) 113-147.

[5] Becker, S. O. and Mathias Hoffmann (2006). "Intra- and International Risk Sharing in the Short-Run and the Long-Run," European Economic Review 50, 777-806.

[6] Botev, Z.I., J. F. Grotowski and D. P. Kroese, "Kernel Density Estimation Via Diffusion" Annals of Statistics, 2010.

[7] Canova, Fabio and Morten Ravn (1996). "International Consumption Risk Sharing," International Economic Review 37, No. 3, August, 573-601.

[8] Cochrane, John (1991) "A Simple Test of Consumption Insurance," Journal of Political Economy 99:5 (October 1991), 957-976.

[9] den Den Haan, Wouter and Andrew Levin (1997) A practitioner's guide to robust covariance matrix estimation, in Handbook of Statistics 15, G.S. Maddala and C.R. Rao, eds., Elsevier (Amsterdam), pp.299-342.

[10] den Haan, Wouter J. & Andrew T. Levin (2000). "Robust Covariance Matrix Estimation with Data-Dependent VAR Prewhitening Order," NBER Technical Working Papers 0255, National Bureau of Economic Research.

[11] Flood, Robert; Nancy Marion and Akito Matsumoto (2009) "International Risk Sharing During the Globalization Era," IMF Working Paper 09/209.

[12] Kose, M. Ayhan, Eswar Prasad, and Marco Terrones (2007) "How Does Financial Globalization Affect Risk Sharing? Patterns and Channels", IMF Working Paper 07/238.

[13] Lewis, Karen (1996) “What Can Explain the Apparent Lack of Consumption Risk-Sharing?” Journal of Political Economy 104:2 (April 1996), 267-297.

[14] Lewis, Karen (1999) “Trying to Explain Home Bias in Equities and Consumption,” Journal of Economic Literature 37:2 (June 1999), 571-608.

[15] Mace, Barbara (1991) “Full Insurance in the Presence of Aggregate Uncertainty,” Journal of Political Economy 99 (October 1991), 928-956.

[16] Mann, H. B.; Whitney, D. R. (1947). "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other". Annals of Mathematical Statistics 18 (1): 50–60.

[17] Obstfeld, Maurice (1994) "Are Industrial Country Country Risks Globally Diversified?" in L. Leiderman and A. Razin, Eds., Capital Mobility: The Impact on Consumption, Investment, and Growth, Cambridge University Press, 1994.

[18] Wilcoxon, F. (1945). "Individual comparisons by ranking methods". Biometrics Bulletin 1 (6): 80–83.

Table 1: Correlations with Largest Trading Partner

* indicates that correlation with trading partner is significantly higher than correlations with other countries using Whitney-Mann U-test, 10% critical value

		k: Differencing Horizon						
Region/ Country	**Largest trading partner**	**1**	**4**	**8**	**12**	**16**	**20**	**24**
Europe								
Austria	Germany	*	*	*	*	*	*	*
Belgium	Germany	*					*	*
Finland	Germany							
France	Germany							
UK	Germany							
Italy	Germany							
Luxembourg	Germany	*	*	*	*	*	*	*
Netherlands	Germany	*	*	*	*	*	*	*
Sweden	Germany							
Spain	France	*		*	*	*	*	*
Germany	France							
Ireland	UK		*	*	*	*	*	*
Norway	UK							
Iceland	Netherlands							
Portugal	Spain			*	*	*		
Australasia								
Australia	Japan							
Japan	US							
New Zealand	Australia	*	*	*	*	*	*	
North America								
Canada	US	*	*	*				
Mexico	US	*	*					
USA	Canada	*	*	*	*	*	*	*

Note: This test does not account for the serial correlation in the k-differences used to construct the correlation coefficients.

Table 2: Sub-Sample Test for No Change in Correlation

Country	Largest trading partner	k: Differencing Horizon						
		1	4	8	12	16	20	24
Australia	Japan							+
Austria	Germany	-	-	-				
Belgium	Germany	-	-					
Spain	France	-	-					
Finland	Germany	-						
France	Germany	-	-	-				
UK	Germany							
Germany	France	-	-	-				
Ireland	UK			+		+	+	+
Iceland	Netherlands	-						
Italy	Germany							
Japan	US							
Luxembourg	Germany							
Mexico	US							
Netherlands	Germany	-	-					
Norway	UK	-						
New Zealand	Australia							
Portugal	Spain	-	-					
Sweden	Germany					+		
USA	Canada							

Notes:

" - " : correlation with trading partner significantly lower in first sub-period: 1960:1-1984:4

" +" : correlation with trading partner significantly lower in second sub-period: 1985:1-2009:2

Tests performed on z-transform of correlation coefficient using VAR-HAC standard errors.

Figure 1: Consumption Correlations

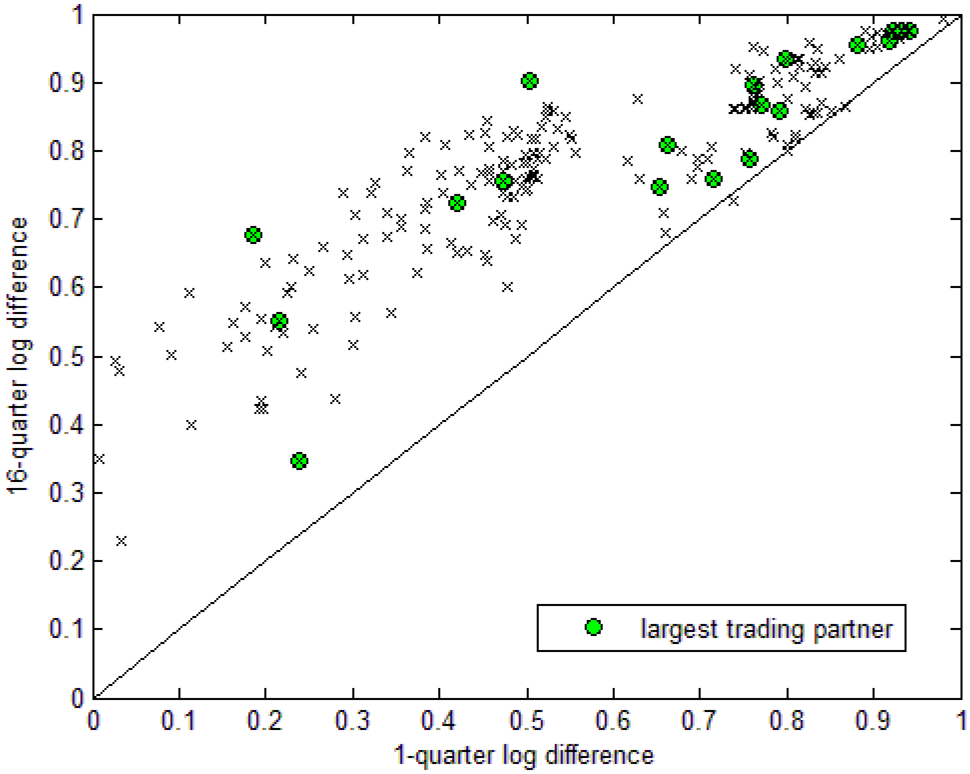


Figure 2: Kernel Densities of Consumption Correlations

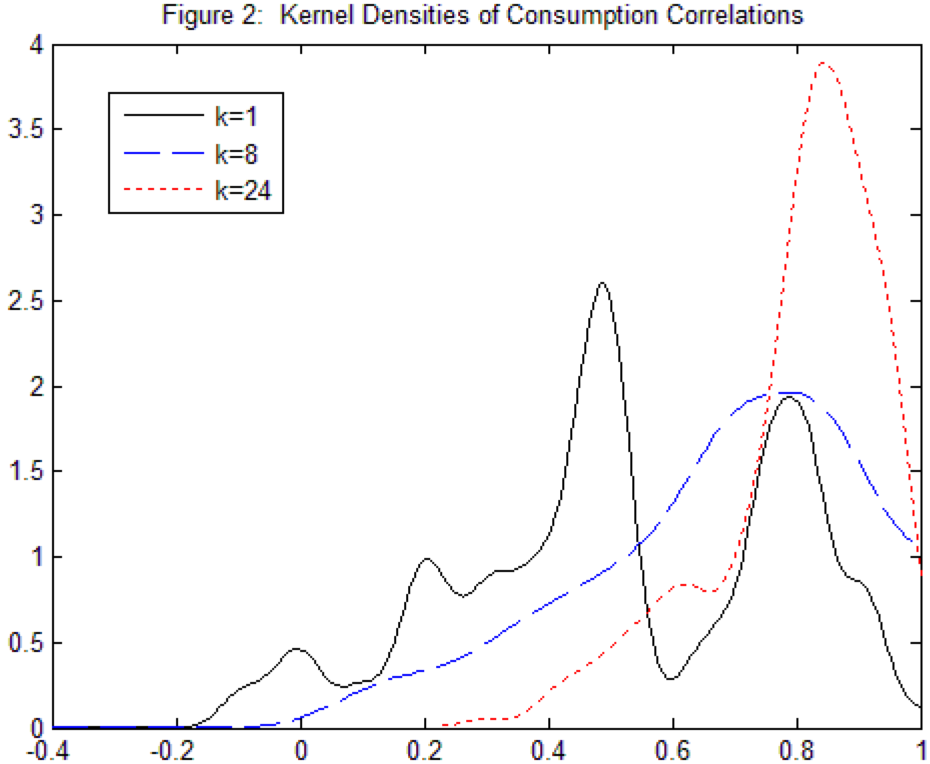
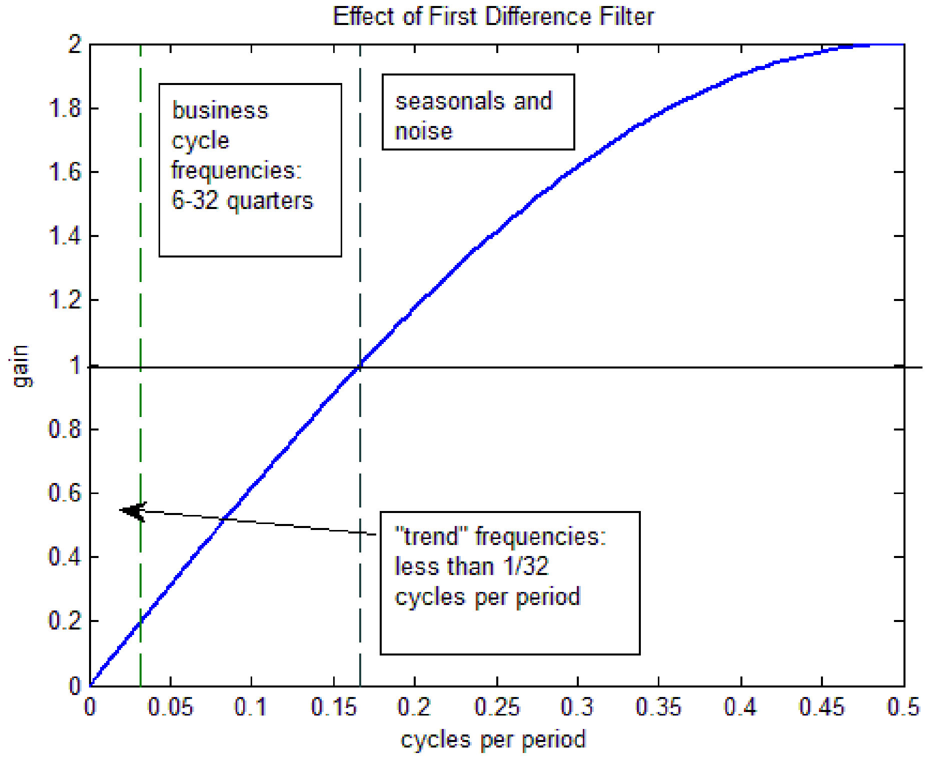


Figure 3



Effect of First Difference Filter
gain
cycles per period
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
business cycle frequencies: 6-32 quarters
seasonals and noise
"trend" frequencies: less than 1/32 cycles per period

Figure 4: The effect of measurement error

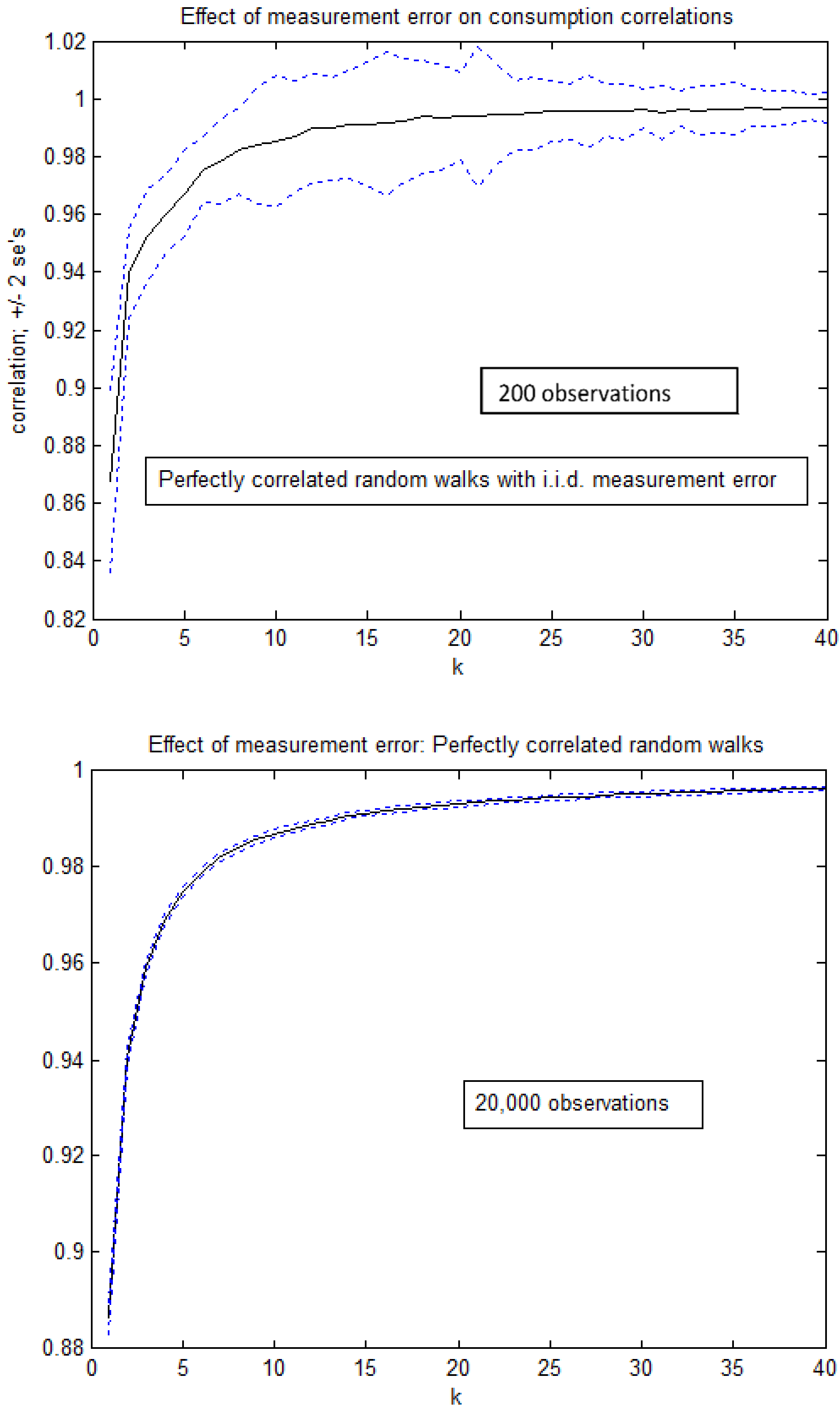


Figure 5: Correlation with major trading partner (solid line) and world (dashed line)

(99% confidence intervals indicated by dotted lines)

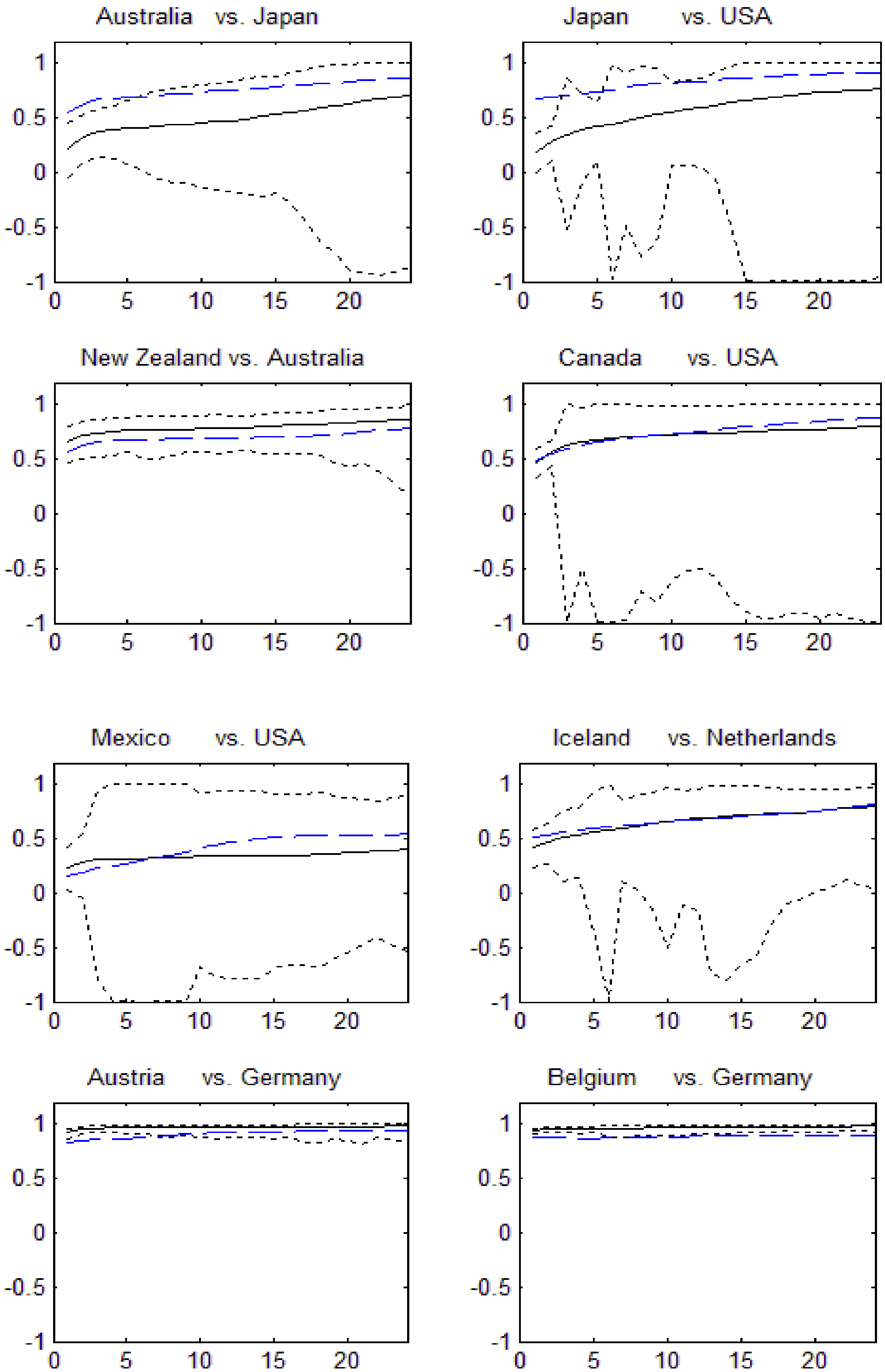


Figure 5, cont'd.

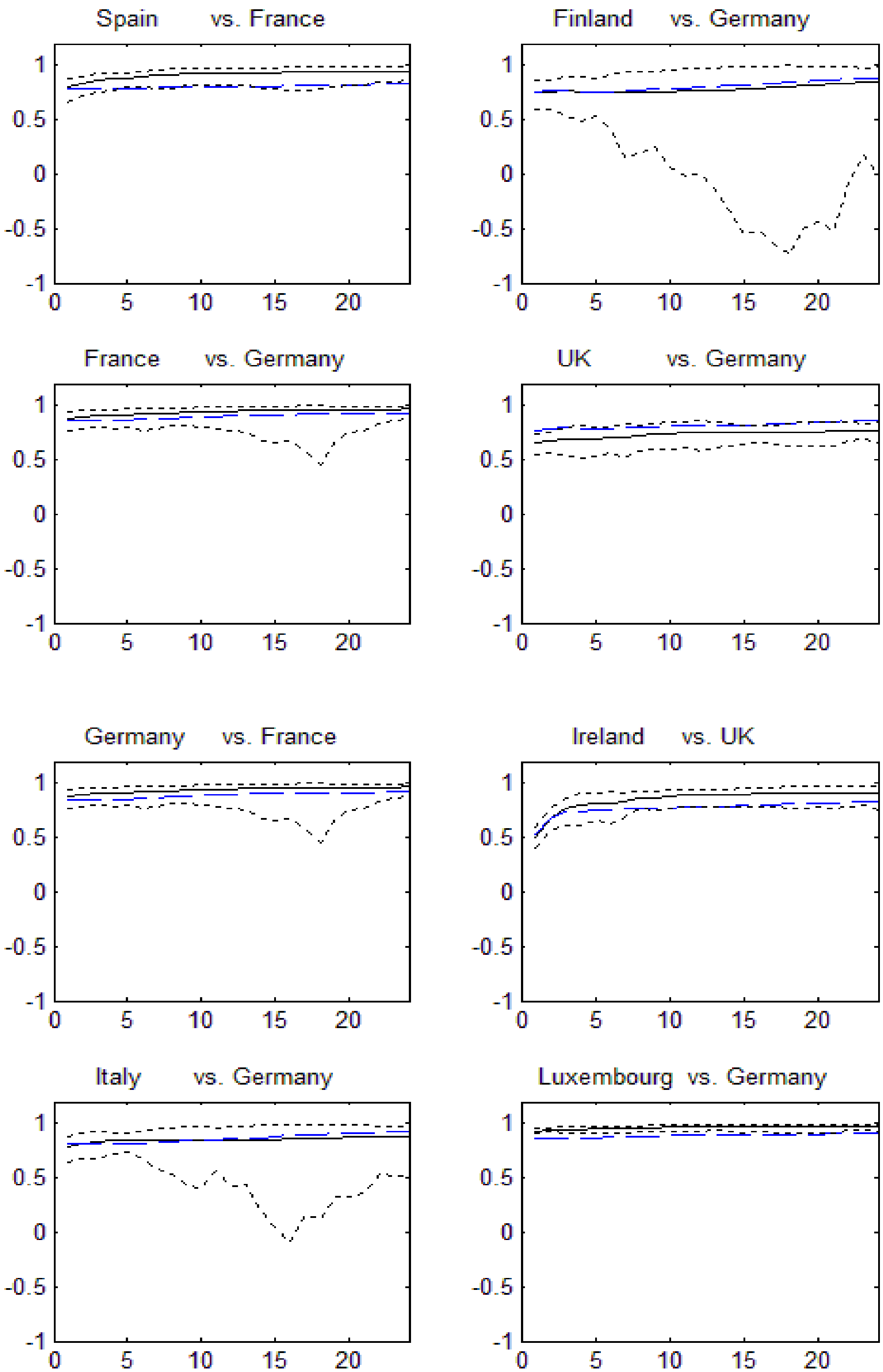
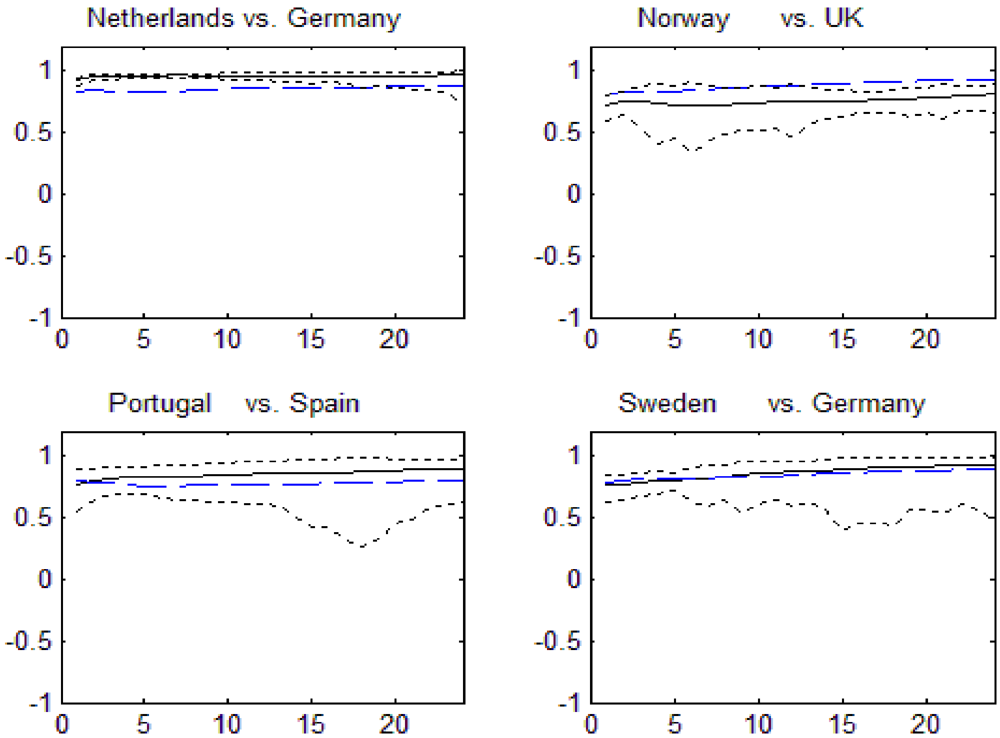


Figure 5 cont'd.



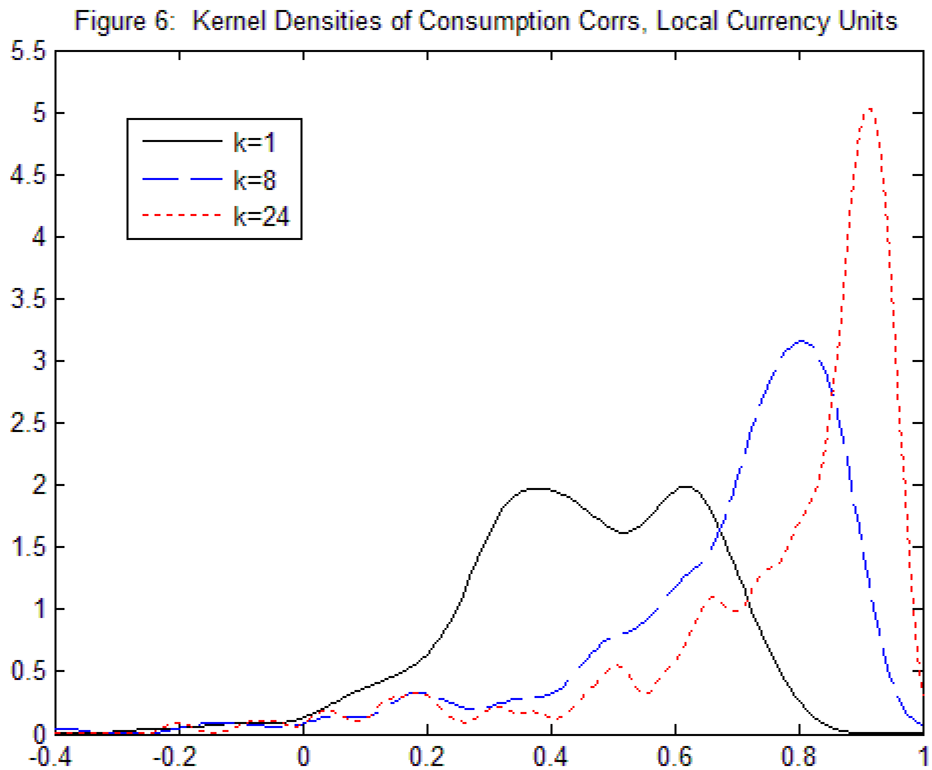
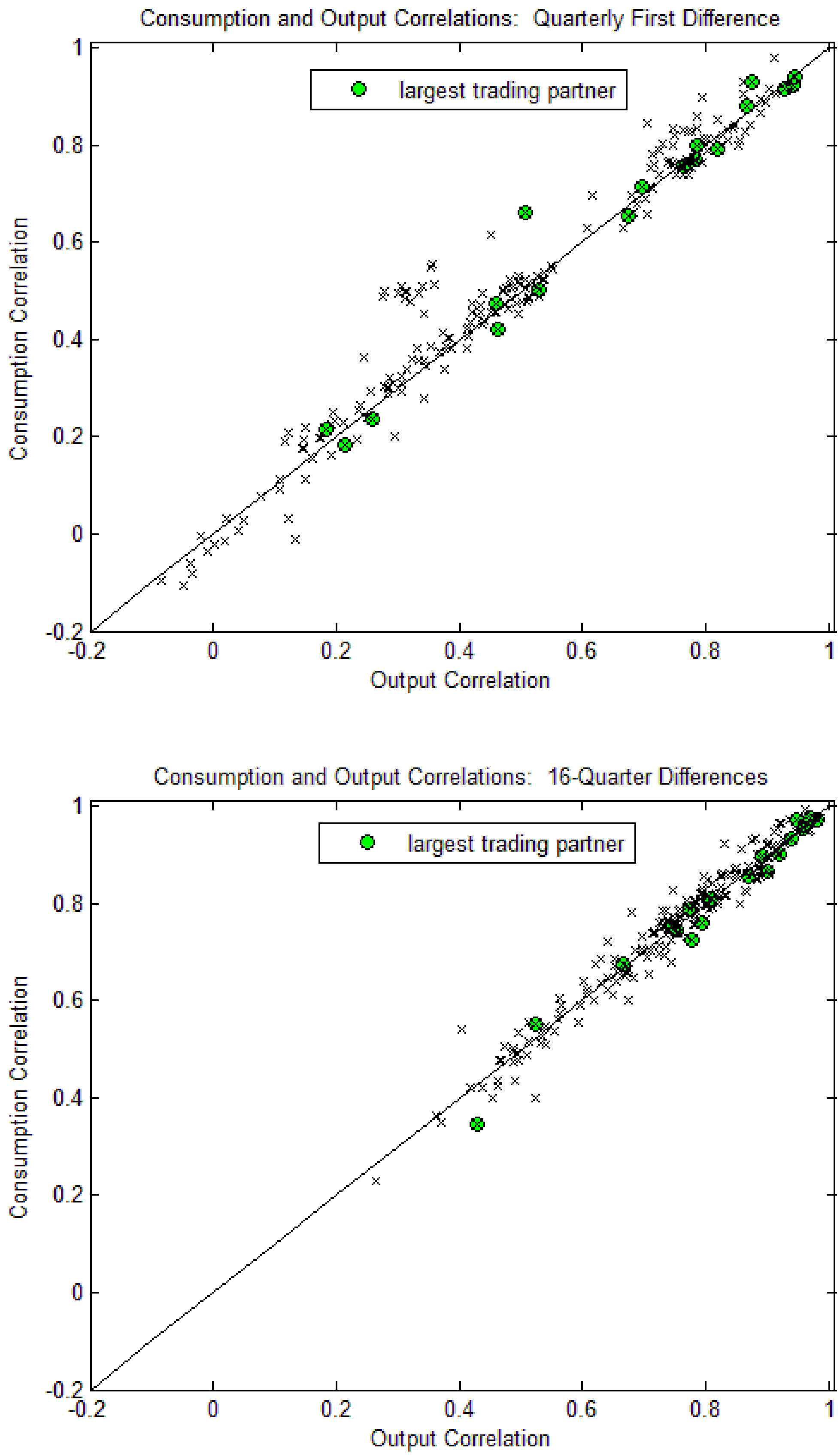


Figure 6: Kernel Densities of Consumption Corrs, Local Currency Units
5.5
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
k=1
k=8
k=24

Figure 7


Consumption and Output Correlations: Quarterly First Difference
largest trading partner
Consumption Correlation
Output Correlation
1
0.8
0.6
0.4
0.2
0
-0.2
-0.2
0
0.2
0.4
0.6
0.8
1
Consumption and Output Correlations: 16-Quarter Differences
largest trading partner
Consumption Correlation
Output Correlation
1
0.8
0.6
0.4
0.2
0
-0.2
-0.2
0
0.2
0.4
0.6
0.8
1