DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS
DIVISION OF AGRICULTURE AND NATURAL RESOURCES
UNIVERSITY OF CALIFORNIA AT BERKELEY

WORKING PAPER NO. 828

ECONOMIC DISTANCE, SPILLOVERS, AND CROSS-COUNTRY COMPARISONS

by

Tim Conley and Ethan Ligon*

*Department of Economics, Northwestern University; and Department of Agricultural and Resource Economics, University of California Berkeley, respectively

California Agricultural Experiment Station
Giannini Foundation of Agricultural Economics
February 1998

ECONOMIC DISTANCE, SPILLOVERS, AND CROSS-COUNTRY COMPARISONS

TIM CONLEY AND ETHAN LIGON

1. Introduction

Popular discussions of economic growth and development often focus on the economic interdependence of nations. Such discussions generally presuppose that the world is connected, and that the economic experience of one nation may not be independent of the experiences of others. Economic theory also suggests that economic outcomes across nations will not be independent; the very existence of international trade lends considerable empirical support to this claim.

From the econometrician's point of view, economic interdependence across countries results in dependence among individual countries' variables. However, the vast majority of empirical work using cross-country comparisons assumes that observations are independent across countries.[1] Such an independence assumption is clearly at odds with commonly expressed beliefs regarding the integration of economies. Moreover, it is often inconsistent with the very economic models being tested with cross-country regressions; for example, those that investigate the impact of international trade on growth.

In this paper, we present empirical methods for cross-country comparisons that explicitly allow for interdependence among countries based upon the econometric model of Conley (1996). This model adds structure to cross-sectional data by using information on *economic distances* between countries. Specifically, countries are modeled as being located in a metric space where the dependence between countries' random variables is a function of the economic distance between them. For example, the covariance between two countries' growth rates is a function of the distance which separates the two countries in this space. Close countries are allowed to have highly correlated variables while variables from countries far away from each other are uncorrelated. This

Date: October 1997.

[1] Notable exceptions include DeLong and Summers (1991),Quah (1992), Quah (1993), Elliott (1993).

modeling strategy provides a simple characterization of possible interdependence among countries, and allows more complicated dependence than group effect or equicorrelation models. Furthermore, it enables us to use estimators that are robust to measurement errors in economic distances. Thus, our modelling strategy is amenable to any application where the econometrician has access to an approximate measure of the economic distance between countries.

We present our empirical methods in the context of a particular application: the study of economic growth across countries. This choice is motivated by the popularity and importance of this topic. We use our spatial model to present empirical evidence in three ways. First, we present a characterization of the correlation structure of growth rates and related economic variables across countries. Second, we estimate an example cross-country regression using our model to allow for dependence in cross-sectional regressions. Finally, we investigate the potential for spillovers to explain growth across countries by looking at whether neighboring countries' characteristics are helpful in predicting growth rates.

The key ingredient in the analysis is the measure of economic distance. A natural first candidate for economic distance is geographic distance. Countries which are physically close—like the U.S. and Canada, or the countries of western Europe—are obviously closely linked, and so geographic distance may be a good measure of economic distance. DeLong and Summers (1991) informally investigate the correlation in growth regression residuals in physically nearby countries but do not find it significant. Elliott (1993) uses geographic proximity in a more formal test of dependence of growth regression errors across countries and finds that spatial dependence is significant. These differences in results may be due in part to differences in approach, but may also arise because geographic distance is poor measure of economic distance for this application. For example, the U.S. may be 'farther' from Mexico than it is from Canada in our economic sense, while Hong Kong and Britain may be rather close. We present results using geographic distance as measurements our economic distance, for comparison to this previous work. However, we construct two alternative measures that, a priori, we think are better measures of economic distance than geographic distance.

We argue that what matters in determining how closely related two economies are has to do with to what degree two countries share common markets, and that this, in turn, has principally to do with the costs of transporting the various factors of production between them. We will concentrate on the factors of physical and human capital and use measures of the cost of transporting these factors as measures of economic distance. The costs of transporting physical capital will have principally to do with the cost of simply moving it, along with trade barriers. Our measurements of this economic distance metric are derived from UPS shipping rates for large packages, which capture some (though certainly not all) trade barriers. The costs of moving embodied human capital involve various barriers to immigration as well as the price of transportation. Our measurements of this economic distance metric are derived from airline fares between countries—a proxy for the cost of transporting an engineer or consultant, for example.

We present parallel results for each of these three economic distance metrics. For each metric, we find evidence of considerable dependence in growth rates of per capita GDP, as well as in a variety of other economic variables. This dependence affects inference in our cross-country regression. In order to better understand the observed spatial dependence of growth rates, we test for spillovers by testing whether near neighbors influence growth. Our results suggest that spillovers are quite important—growth rates in a country depend importantly on observable features of its neighboring countries.

2. Measuring Economic Distance

We identify economic distance in this application with the costs of trade between countries. Here, an economic distance needs to characterize the relationship between countries' growth rates. In equilibrium, factor prices would be expected to be close in countries who had small transaction costs between them. With similar prices, such countries would have correlated output growth and other economic outcomes. Conversely, economic outcomes in countries with very high transaction costs between them who would be nearly independent. Thus measures of transaction costs are relevant to cross-country patterns in growth rates. In reality, such trade costs will certainly vary across different goods. However, to keep the model tractable, we posit a single cost of transacting between countries.

We assume that our measured economic distances define a valid metric and so must satisfy several conditions. The first of these is simply a condition that it cannot be costly to not trade. The second is a condition that there be a positive cost involved in trade. The third is a requirement that it cannot be more costly to ship between two countries than it is to ship via an intermediate nation. Finally, a requirement of symmetry implies that it should be equally costly to ship from country i to j as to ship in the opposite direction. Asymmetric tariff barriers are perhaps the most obvious feature of the world economy which violates this condition. However, we maintain the assumption for simplicity.[2]

The we address the how to measure a 'general' transaction cost for factors by considering two broadly defined factors: physical and human capital. However, it seems apparent that economic distance might be very different for physical capital than for human capital, even if we consider only the costs of transportation of either of these factors. Thus, rather than trying to arrive at a single measure of economic distance, we collect data on several different measures: geographic distance (relevant for both physical and human capital transport); the cost of shipping some benchmark package (suited to physical capital); and the cost of airfare (relevant for embodied human capital). The most appropriate of these metrics will depend on the exact model of growth that is to be evaluated in light of our empirical results. [3]

One issue which concerns us is the possibly endogeneity of some measures of economic distance. In particular, the construction of transportation infrastructure will certainly effect the costs of shipping goods. While it may be possible to collect data on economic distances which are predetermined, if not strictly exogeneous, this task is beyond the scope of this paper.

An important consideration in collecting data on economic distances has to do with the units in which distances are measured. For geographic distance, this is no problem; we simply compute distances in statute miles. For distance measures which try to directly measure the cost of transport, however, prices quoted in nominal units raise concerns about relevant exchange rates and purchasing power parity. We avoid this problem by exploiting our assumed symmetry of economic distance. Suppose that it costs a yen to transport a particular good from Japan to the U.S., and b dollars to transport the good in the other direction. We want to express all costs in terms of U.S. dollars, and so we use b/a as the relative "exchange rate" between Japan and the U.S. Similar exchange rates are computed for all other countries, again using the U.S. as a benchmark.

2.0.1. *Geographic Distance.* The first measure of economic distance we propose is the most obvious; simple geographic distances. To measure

[2]Relaxing symmetry would mean that we couldn't work with a metric space. While some of our estimators could be adapted to this case, we're unsure as to how we might go about interpreting the results.

[3]The data and code we describe below are available from `ligon@are.berkeley.edu`.

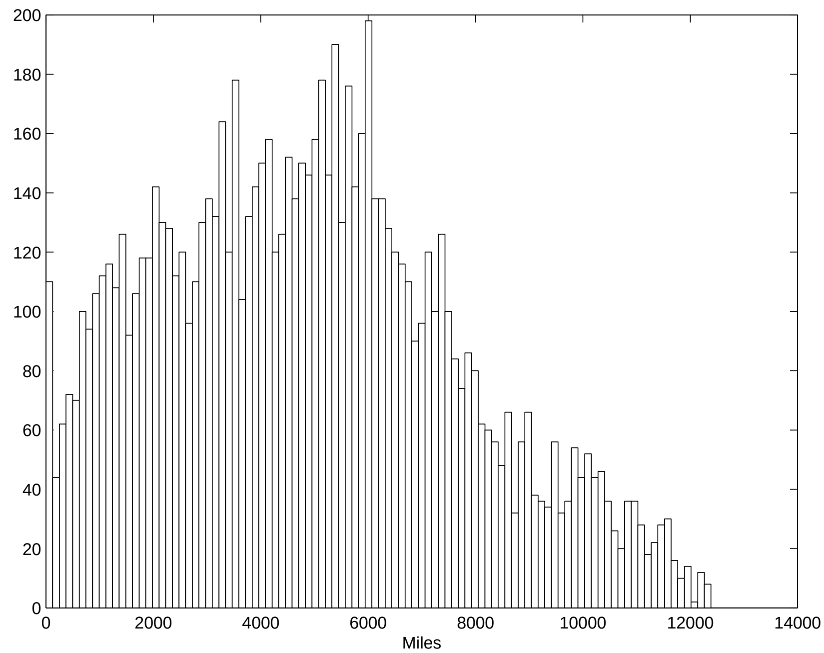

FIGURE 1. Histogram of Geographic Distances

geographic distances between countries, we collected data on the latitude and longitude of the capital cities of each of the nations represented in the Penn World Tables (Summers and Heston 1991). Assuming a spherical planet, we then use these data to compute the "great circle" distance between capitals.[4] Figure 1 is a histogram of these geographic distances, while Table 1 provides some examples of geographic distances in its first panel.

2.0.2. *UPS Distance.* As a way of measuring the costs of transporting physical capital, we asked the United Parcel Service to provide us with the cost of shipping a 20 kilogram express package between capital cities.[5] Consistent with our earlier claim that economic distance

[4]These measures are inexact—the circumference of the earth about the equator is approximately 42 miles greater than the circumference about the poles. By averaging these two circumferences, the error in our distances can be no greater than 0.1 per cent, neglecting surface features of the planet.

[5]The nomenclature for the class of shipping we wanted actually varies somewhat by the nation of origin. In the U.S., such packages are delivered overnight, and are called "Express;" this service is called other things in other places, and is not, for the most part, overnight. Where overnight service was unavailable, we took whatever service promised delivery of parcels (rather than letters) in the fewest number of days (typically two to three).

Geographic Distance (Miles)

	Australia	Brazil	Egypt	France	India	S. Africa	UK	US
Australia	0	8732	8870	10520	6438	6696	10558	9906
Austria	9895	5927	1484	642	3459	5680	766	4429
Brazil	8732	0	6143	5428	8852	4271	5470	4223
Colombia	8966	2287	6968	5362	9452	6547	5281	2367
Denmark	9971	6055	1994	638	3636	6202	593	4049
Egypt	8870	6143	0	1997	2752	4500	2182	5814
Finland	9458	6602	2102	1187	3244	6513	1133	4312
France	10520	5428	1997	0	4095	5807	212	3834
Germany	9990	5966	1797	545	3594	5981	578	4174
India	6438	8852	2752	4095	0	5782	4171	7492
Ireland	10717	5383	2470	483	4399	6209	288	3386
Israel	8699	6404	264	2073	2501	4670	2244	5906
Portugal	11224	4529	2360	903	4833	5321	985	3568
S. Africa	6696	4271	4500	5807	5782	0	6010	7900
Taiwan	4547	11507	5388	6107	2728	7878	6079	7864
Thailand	4650	10329	4518	5871	1814	6307	5925	8802
UK	10558	5470	2182	212	4171	6010	0	3669
US	9906	4223	5814	3834	7492	7900	3669	0

UPS Distance (Dollars)

	Australia	Brazil	Egypt	France	India	S. Africa	UK	US
Australia	0	427	309	267	245	309	217	224
Austria	249	430	352	246	210	280	208	192
Brazil	427	0	457	474	475	534	396	397
Colombia	309	397	421	361	421	421	352	397
Denmark	270	479	323	246	323	323	208	192
Egypt	309	457	0	361	421	421	291	397
Finland	270	479	323	246	323	323	208	192
France	267	474	361	0	351	361	157	176
Germany	267	367	350	106	344	350	106	171
India	245	475	421	351	0	421	238	354
Ireland	270	479	361	246	361	361	125	192
Israel	389	655	500	332	475	500	259	397
Portugal	300	436	361	246	361	361	208	258
S. Africa	309	534	421	361	421	0	291	402
Taiwan	182	419	315	193	182	315	193	172
Thailand	187	419	394	226	317	394	215	224
UK	217	396	291	157	238	291	0	198
US	224	397	397	176	354	402	198	0

Air Fare Distance (Dollars)

	Australia	Brazil	Egypt	France	India	S. Africa	UK	US
Australia	0	2478	2227	2257	2141	3568	2151	1853
Austria	2541	1661	1464	754	1808	2362	945	688
Brazil	2478	0	1617	907	1845	2207	790	987
Colombia	2688	743	2044	1334	2198	2560	1143	1197
Denmark	2553	1369	1200	665	1634	1996	579	700
Egypt	2227	1617	0	710	980	2318	901	1100
Finland	2713	1653	1716	1006	1918	2280	863	860
France	2257	907	710	0	1054	1608	191	670
Germany	2427	1372	678	506	1405	1999	582	890
India	2141	1845	980	1054	0	2472	1055	1135
Ireland	2321	960	1027	317	1225	1586	170	618
Israel	2220	1560	288	760	766	2187	770	1200
Portugal	2504	1143	973	263	1317	1770	353	738
S. Africa	3568	2207	2318	1608	2472	0	1417	1913
Taiwan	2594	1397	367	490	1347	2098	681	1160
Thailand	2151	1397	950	490	947	2098	681	980
UK	2151	790	901	191	1055	1417	0	497
US	1853	987	1100	670	1135	1913	497	0

TABLE 1. Examples of distances between countries' capitals.

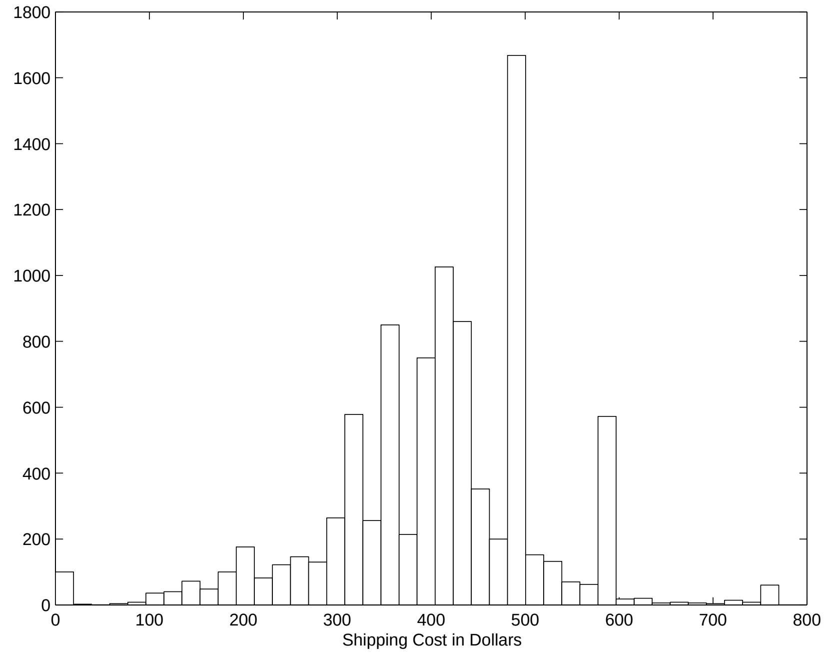

FIGURE 2. Histogram of UPS Distances

might vary according to commodity, UPS rates often vary according to the content of the package. Where it mattered, we specified that the package was a document, and that rates should be calculated F.O.B.

We should add that we were not able to obtain UPS data for each pair of capital cities in the Penn World Tables.[6] Instead, we got complete fare information for the 28 selected countries ('hubs') which seemed to be the most likely trans-shipment points, and which also had adequate data.[7] Where possible, we used rates current as of November 1988, so as to avoid the complications which attended the breakup of the Soviet Union.

UPS distances are computed using the minimum cost path through the network of hubs. This leaves us with a nearly complete set of distances.[8] Many of these distances are computed using the triangle

[6]As there are 139 countries, 19182 inquiries of UPS would perhaps have been too much of an imposition.

[7]These are Australia, Austria, Belgium, Brazil, Chile, China, Colombia, Denmark, Egypt, Finland, France, Germany, Guatemala, India, Ireland, Israel, Korea, Norway, the Phillipines, Portugal, South Africa, Spain, Sweden, Taiwan, Thailand, the United Kingdom, the United States, and the Soviet Union.

[8]Complete save for distances to and from Bhutan, which could not be reached from any of the UPS hubs we used.

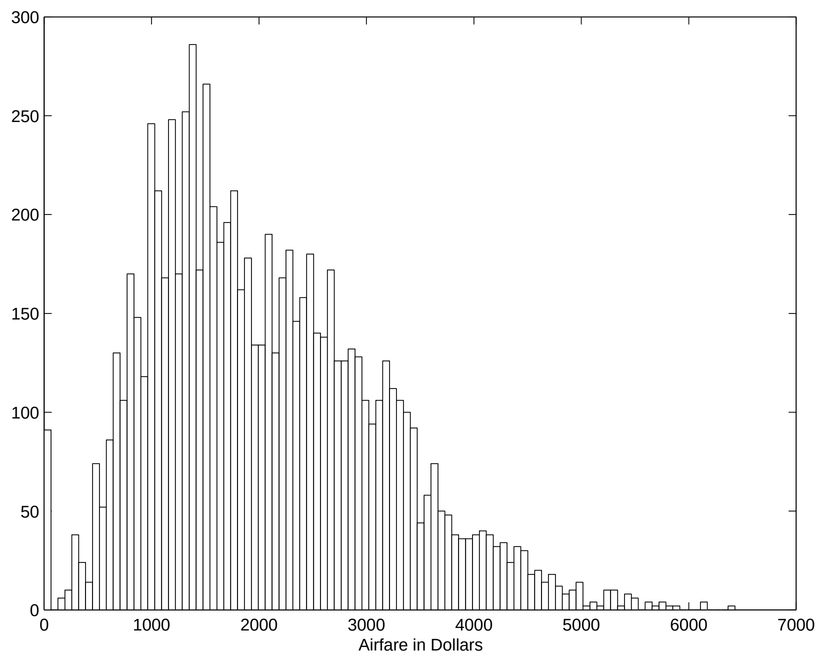

FIGURE 3. Histogram of Airfare Distances

inequality. While in principle these computed distances are merely upper bounds on the true distances, in practice UPS shipping usually involves an intermediate hub. Figure 2 is a histogram of the computed UPS distances, to give some idea of the distribution of distances; Table 1 gives a few examples in its second panel.

2.0.3. *Airfare Distance.* While UPS distances were intended to capture something of the costs of transporting physical capital, we would also like to come up with some measure of the costs of transporting embodied human capital. The measure that we use is the cost of airfare between countries' capitals.

Our calculation of airfares was quite similar to our calculation of UPS distances. Our data is from SABRE, an airline reservation system used by travel agents. We chose 15 hub cities.[9] Unlike the UPS case, we did not require the hubs to be capital cities, but rather used hubs that came up repeatedly in our early inquiries. We used the lowest cost fare from the hubs to each of the capital cities of the countries in the Penn World Tables. The date of travel was either February

[9]Buenos Aires, Brasilia, London, Paris, Jeddah, Jerusalem, Singapore, Hong Kong, Cairo, Tokyo, Washington D.C., Mexico City, Riyadh, New York, Sydney, Los Angeles.

1, 1997, or the closest available day thereafter; all fares were booked roughly six months in advance, the fare class was coach with restrictions (where applicable). As with the UPS distances, we used the minimum cost means of travelling through the hub network. Figure 3 gives a histogram of the cost of airfares, while the third panel of Table 1 gives some example airfares.

3. Econometric Model

This section contains a brief description of the model of data generation. A rigorous exposition of this model for the special case where economic distances are Euclidean can be found in Conley (1996). The model is that individual countries reside at locations in the metric space (Λ, d), with each individual i located at a distinct point $s_i \in \Lambda$. Random variables associated with each country are indexed by position s_i and are called random fields. The econometrician's sample consists of a observations at points within one of a sequence of regions $\{\Lambda_N\}$ with $\Lambda_N \subset \Lambda_{N+1}$. Within Λ_N, the econometrician observes T_N realizations of random fields at a collection of locations $\{s_i\}_{i=1}^{T_N}$ inside Λ_N, one corresponding to each country. The econometrician also observes measurements of distances between locations $\{d(s_i, s_j)\}$. We will take limits letting sample regions Λ_N increase (as $N \to \infty$).

The basic model of dependence is that the distance between countries' positions, corresponding to their economic distances, characterizes the dependence between their random fields. If two locations s_i and s_j are close then their random fields, say X_{s_i} and X_{s_j}, may be very highly correlated. As the distance between s_i and s_j grows large, the random fields X_{s_i} and X_{s_j} become closer to being independent.

More formally, we assume that the random field X_s is stationary, mixing, and isotropic.[10] The concepts of stationary and mixing for random fields are straightforward generalizations of their time series counterparts. Stationarity simply means that the joint distribution of X_s for any collection of locations $\{s_i\}_{i=1}^{m}$ (i.e., $\{X_{s_1}, X_{s_2}, ..., X_{s_m}\}$) is invariant to shifts in the entire set of locations $\{s_i\}_{i=1}^{m}$. Mixing means that the random fields X_{s_i} and X_{s_j} become asymptotically independent as the distance between s_i and s_j goes to infinity. Our final assumption is that X_s is *isotropic*, this means that the $\text{cov}(X_{s_i}, X_{s_j})$ depends only on the distance between s_i and s_j, not the direction.

[10]The assumption of stationarity could be relaxed. For example, moments could be allowed to vary with location, provided their spatial averages converged. The estimators we describe below could then be interpreted as spatial averages of these space-varying moments.

3.1. **Spatial Correlation Estimation.** The first question we want to investigate is whether there is significant spatial dependence in countries' variables using a given economic distance metric. For simplicity we will focus on just one of the implications of spatial independence, that of zero spatial correlation: $\text{cov}(X_{s_i}, X_{s_j}) = 0$ for $s_i \neq s_j$. Since the random field X_s is stationary and isotropic, $\text{cov}(X_{s_i}, X_{s_i})$ is simply a function of $d(s_i, s_j)$, and can be represented in a simple graph.

To construct such a graph, we use a nonparametric estimator of the spatial autocovariance function. The estimator is essentially that proposed by Hall, Fisher, and Hoffman (1992). The autocovariance at distance k is estimated by a local average of cross-products of observations that are close to k units apart,

$$\widehat{C}(k) = \sum_{i=1}^{T_N} \sum_{j=1}^{T_N} W\left(k, d(s_i, s_j)\right)(X_{s_i} - \bar{X})(X_{s_j} - \bar{X}).$$

Here $\bar{X}$ is the sample average of X_s. The weighting function W is normalized so that the weights sum to one. We also require W to be a function of sample size that will concentrate its mass at zero as the sample becomes arbitrarily large. Thus, in large samples, the covariance at lag k will be estimated by an average of cross-products of only those observations that are very close to k units apart. This estimator can also be viewed as a non-parametric regression of cross-products $(X_{s_i} - \bar{X})(X_{s_j} - \bar{X})$ on the distances between s_i and s_j.

We take a slightly unusual approach to conducting a pointwise test of whether our estimate of $\widehat{C}(k)$ consistent with spatial independence. Instead of calculating pointwise standard error bands for our estimates, we plot an acceptance region for the specific null hypothesis of spatial independence. Then our pointwise hypothesis test can be done by simply observing whether our estimate lies inside the acceptance region. We prefer this bootstrap method to pointwise standard error bands because asymptotic standard errors for these local-average estimators will not 'correct' for spatial dependence. This is analogous to lack of corrections for serial dependence in limiting distributions of local-average time series estimators (see e.g. Robinson (1983)). Since *a priori* we expect there to be a considerable amount of dependence between observations we want to avoid overstating the information in our sample by using estimates of pointwise standard errors that abstract from spatial dependence.

To get an acceptance region for the hypothesis of spatial independence we employ a simple bootstrap technique. We hold the sample locations fixed and simulate draws from a distribution with the same stationary (marginal) distribution as our data, but from one which has spatial independence. To do this simulation, we independently sample with replacement from the empirical marginal distribution of our variables. For each of these bootstrap samples, which by construction have zero spatial correlation, we can calculate an estimate of the correlation function exactly as we had done for the original data. For each value of k we take an envelope containing say 95 per cent of our bootstrap estimates to give us a rough acceptance region for the hypothesis of spatial independence. Thus, if our point estimate lies outside of this envelope, there is evidence that there is significant spatial dependence.

If our distances are measured exactly, then under suitable regularity conditions $\widehat{C}(k)$ will be a consistent estimator of the autocovariance function of X_s. If the distances between observations are measured with error then $\widehat{C}(k)$ will recover some unknown weighted average of the true autocovariances. These unknown weights will be determined by the distribution of the measurement error. Despite the fact that the weights are unknown, the fact that they must be positive implies that $\widehat{C}(k)$ will still contain useful information about whether there is spatial dependence in X_s. We can still compute an acceptance region for the statistic $\widehat{C}(k)$ under the null of spatial independence so we can still test this hypothesis. Of course, it suffers a loss of power versus some alternatives relative to using an estimate of the autocovariance function iself.

3.2. **Growth Regression Estimation.** We model the growth rate of country i, Y_{s_i}, as a linear function of explanatory variables X_{s_i} and an additive error term u_{s_i} that is uncorrelated with X_{s_i}, or

$$\begin{aligned} Y_{s_i} &= X'_{s_i}\beta + u_{s_i} \\ EX_{s_i}u_{s_i} &= 0. \end{aligned}$$

We use this moment condition to derive the GMM estimator of β, which we call b_N— in this case of course, this coincides with the OLS estimator,

$$b_N \equiv \left[\frac{1}{T_N}\sum_{i=1}^{T_N} X_{s_i}X'_{s_i}\right]^{-1}\left[\frac{1}{T_N}\sum_{i=1}^{T_N} X_{s_i}Y_{s_i}\right]. \tag{1}$$

Under suitable regularity conditions,[11] the asymptotic distribution of b_N is

$$\sqrt{T_N}(b_N - \beta) \Rightarrow N\left(0, \left[Ex_{s_i}x'_{s_i}\right]^{-1} V \left[Ex_{s_i}x'_{s_i}\right]^{-1}\right)$$

where $\Rightarrow$ denotes convergence in distribution and V is defined by

$$V = \lim_{N\to\infty} E\frac{1}{T_N}\left[\sum_{i=1}^{T} X_{s_i}u_{s_i}\right]\left[\sum_{i=1}^{T} X'_{s_i}u_{s_i}\right].$$

The matrix V can be interpreted as the spectral density of the spatial process $X_{s_i}u_{s_i}$ at frequency zero. Thus this result is analogous to the asymptotic covariance matrix for the GMM estimator in the more familiar time series context (Hansen 1982). Conley (1996) provides a class of estimators of V and shows they are consistent for Euclidean economic distance metrics. We use a particularly simple member of that class that is feasible even without Euclidean economic distances:

$$\hat{V} \equiv \frac{1}{T_N}\sum_{d(s_i,s_j)<m} X_{s_i}X'_{s_j}\hat{u}_{s_i}\hat{u}_{s_j}, \tag{2}$$

where $\hat{u}_{s_i}$ is a residual, and$m$ is a cutoff value for inter-point distances $d(s_i, s_j)$. This estimator is analogous to a time series asymptotic covariance matrix estimator that simply adds up the first m sample autocovariances. For $\hat{V}$ to be consistent m must be allowed to grow but at a rate slower than that at which the sample size increases.

4. Results

We use our spatial model to present empirical evidence in three ways. First we will present estimates of the correlation structure across countries of growth rates. For our measure of economic growth, we use growth in per capita GDP over the period 1960–85. We also investigate correlations across countries in a set of economic variables often used to 'explain' differences in growth rates. Next, we estimate an example cross-country regression to determine whether spatial correlation patterns in growth rates can be rationalized by spatial correlation in regressors—commonly viewed as being determinants of growth. In these regressions we allow for dependence across countries as a function of economic distance and find that this impacts inferences drawn

[11] Conley (1996) provides a set of conditions for consistency and asymptotic normality of GMM estimators with this model of spatial dependence for Euclidean economic distance metrics. Among these are that (Y_{s_i}, X'_{s_i}) is a stationary mixing process with slightly more than fourth moments, and the sampling point process determining $\{s_i\}$ is also stationary and mixing.

about regressors significance. Finally, we investigate the potential for spillovers to explain growth rates across countries. We do this by looking at whether neighbouring countries' predicted growth rates are correlated with residuals from our growth regression. In other words, we test whether a countries' neighbors' characteristics would be helpful in predicting its growth rate.

4.1. **Spatial Correlation.** In this section we present estimates of spatial correlation functions for growth rates and their potential determinants. In order to allow for a country specific idiosyncratic component we allow these covariance functions to be discontinuous at zero. To see the connection between idiosyncratic components and covariance function discontinuity consider a variable u_{s_i} that can be written as the sum of two independent variables λ_{s_i} and ε_{s_i} :

$$u_{s_i} = \lambda_{s_i} + \varepsilon_{s_i},$$

where λ_{s_i} is spatially correlated with a continuous covariance function $C_\lambda(\cdot)$, and where ε_{s_i} is an idiosyncratic component that is spatially independent and so has a covariance function $C_\varepsilon(\cdot)$ that is nonzero only at the origin. Denoting the covariance function for u as $C_u(\cdot)$, we write

$$C_u(d) = \begin{cases} C_\lambda(0) + C_\varepsilon(0) & \text{for } d = 0 \\ C_\lambda(d) & \text{for } d > 0. \end{cases}$$

Thus the covariance function of the variable u is discontinuous at the origin. We use the nonparametric estimator described above in Section 3.1 using a normal kernel for the weight function $W()$ to estimate covariances for non-zero distances, and we estimate covariance at distance zero by sample second moments.[12]

Figure 4 displays computed spatial autocorrelation functions using each of the three distance measures discussed in Section 2. The bandwidths in our normal kernel were chosen to be 300 miles for geographic distance, \$50 for UPS distance, and \$150 for our airfare distance.

Each panel of Figure 4 provides strong evidence of spatial autocorrelation between neighboring countries. The solid lines in this figure are the estimated spatial autocorrelations, while the dotted region is a 95 per cent acceptance region for the null hypothesis of spatial independence. Each panel in Figure 4 has distances ranging from 0 to 35 per cent of the largest distance; for geographic distances this is 0

[12]In principle we could estimate the two covariance functions $C_\lambda(\cdot)$ and $C_\varepsilon(\cdot)$, using the information in sample covariances at very small distances to estimate $C_\lambda(0)$. However, in this application there are too few observations in the relevant range for this to be feasible.

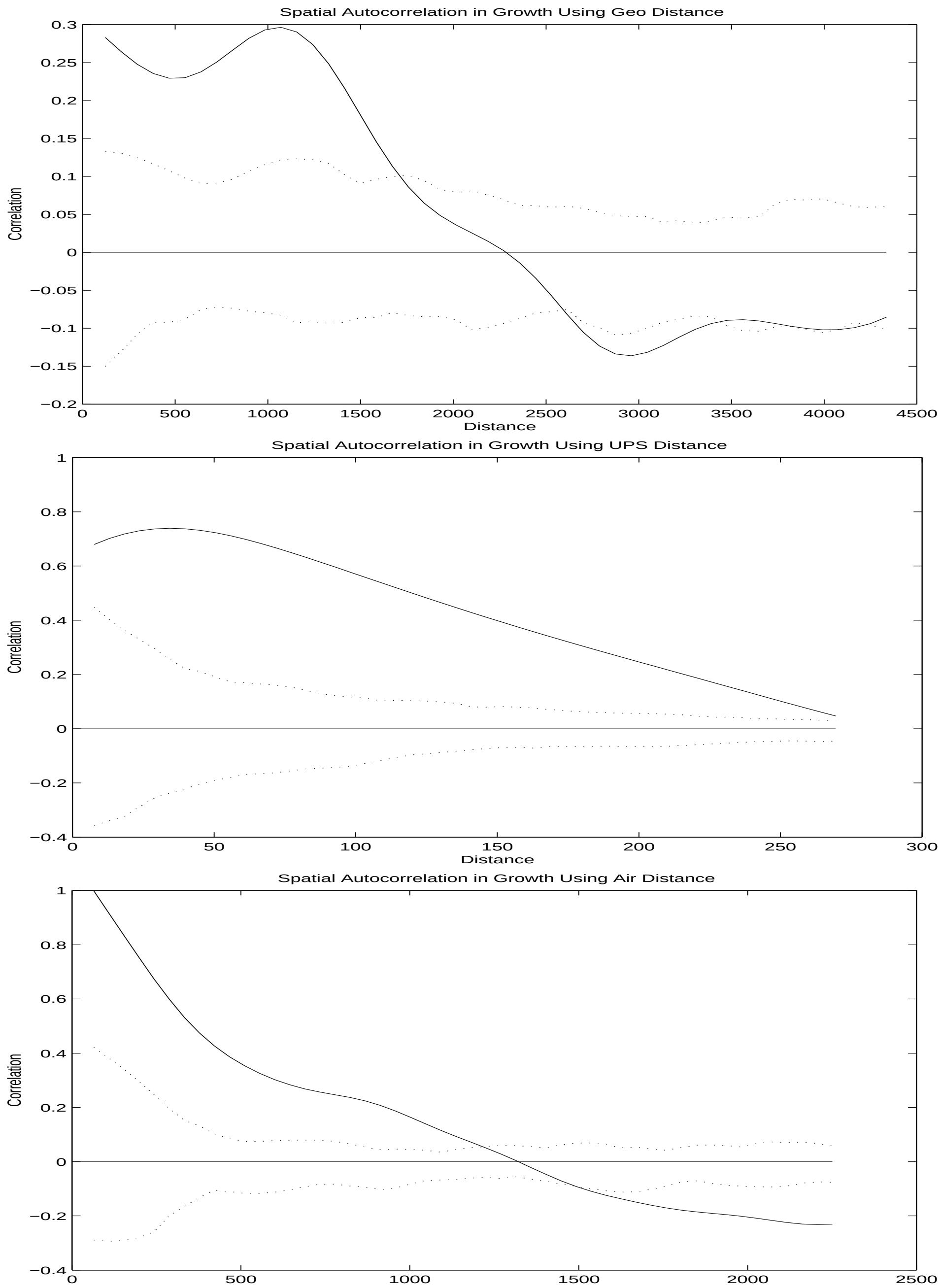

FIGURE 4. Spatial Autocorrelation of per capita GDP growth, 1960–85. The first panel uses geographic distance, the second UPS distance, and the third air fare distance. The dotted regions in each figure are bootstrapped 95 per cent acceptance regions for the null hypothesis of independence.

to about 4400 miles, while UPS and airfare distances range from 0 to about \$260 and \$2250, respectively.

The largest autocorrelations are observed using the UPS and airfare distances; growth rates between nearby neighbors using these measures have a correlation coefficient as high as one.[13] Estimated correlations are somewhat smaller using geographic distance, reaching a maximum of 0.3. While the spatial ACF is basically monotonically declining using airfare distance, it is non-monotone for the other two distances, reaching its maximum at about 1200 miles and \$70 for geographic and UPS distances, respectively. To summarize some of the information in Figure 4, we can reject the hypothesis of independence for countries less than 1700 miles from each other when using geographic distance. Similarly, there appears to be significant positive spatial correlation between the growth rates of countries with UPS distances less than about \$270, or airfares less than about \$1200.

A natural question to ask is whether the observed patterns of correlation are due to correlations in various 'determinants' of growth. For example, it's possible that all of the spatial correlation we observe in growth rates is actually due to spatial variation in levels of per capita GDP in 1960, with otherwise spatially independent increments to GDP since. To investigate this, we simply use a set of variables commonly used in the empirical growth literature (Levine and Renelt 1992).

Figures 5–7 display estimated spatial autocorrelations for a number of variables which have played a prominent role in the empirical growth literature. A complete description of these data may be found in Barro (1991). Briefly, however, they include measures of real per capita GDP in 1960 (GDP60) and 1985 (GDP85), secondary and primary school enrollment rates (SEC60 and PRIM60), the share of real government consumption expenditures to real GDP for 1970–1985 (GOVCONS), revolutions from 1960–1985 (REV), assassinations per million (ASSASS), and the absolute deviation of 1960 PPP investment deflator from its sample mean (PPI60DEV).

These figures reveal that in fact several of these variables do exhibit a signicant spatial autocorrelation. Particularly notable for their high degree of spatial dependence are GDP, measured in both 1960 and 1985, and the education variables SEC60 and PRIM60. It's interesting to note that the spatial correlation in GDP is greater in 1985 than in 1960; for small to moderate distances, the spatial ACF for GDP85 lies

[13] The estimator of spatial autocorrelations presented in Section 3 is not constrained to produce estimates in the interval $[-1, 1]$. Estimates presented in Figure 4 and elsewhere impose this constraint.

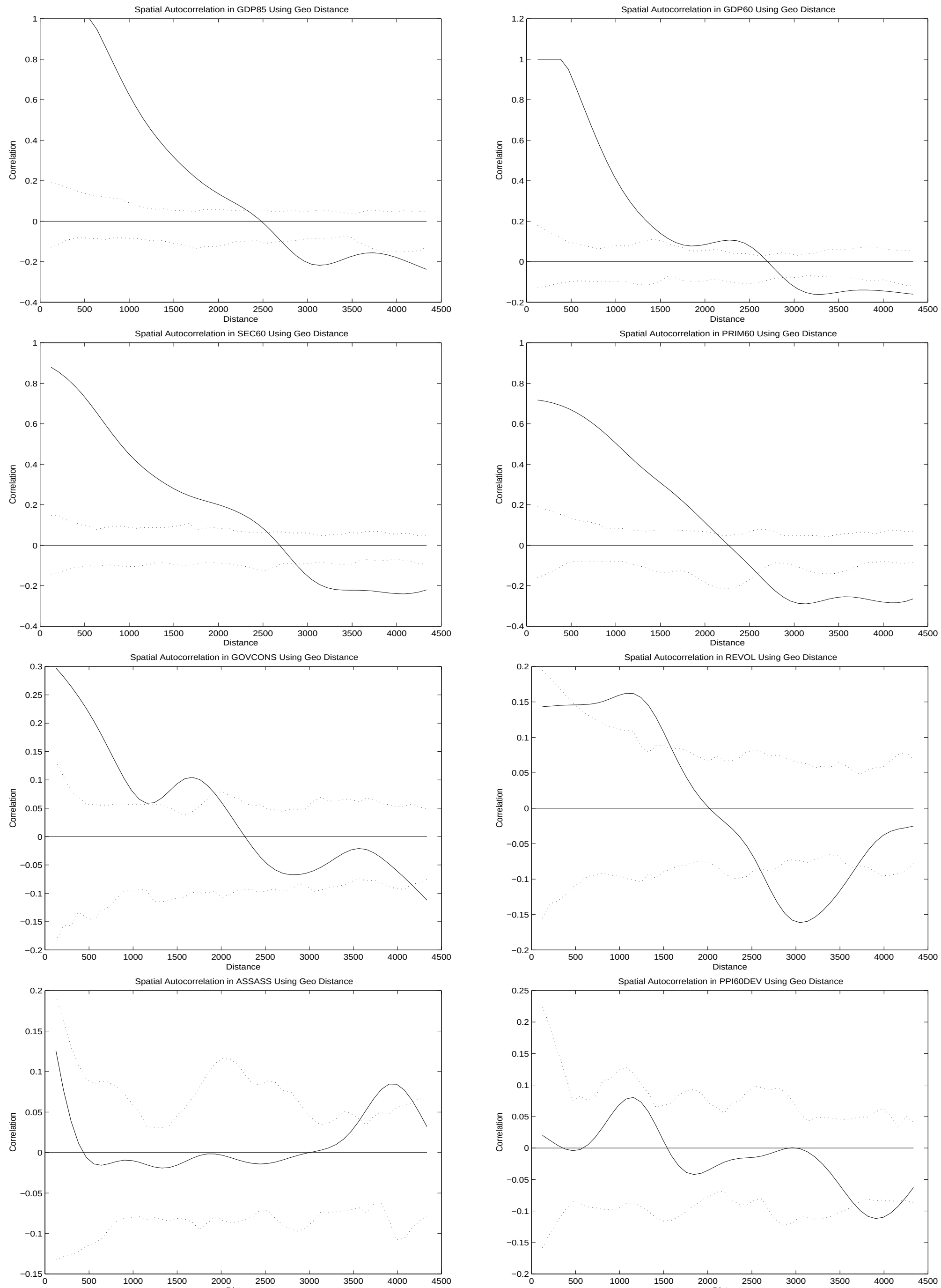

FIGURE 5. Spatial Autocorrelation of Various Variables, using Geographic Distance. The dotted regions in each figure are bootstrapped 95 per cent acceptance regions for the null hypothesis of independence.

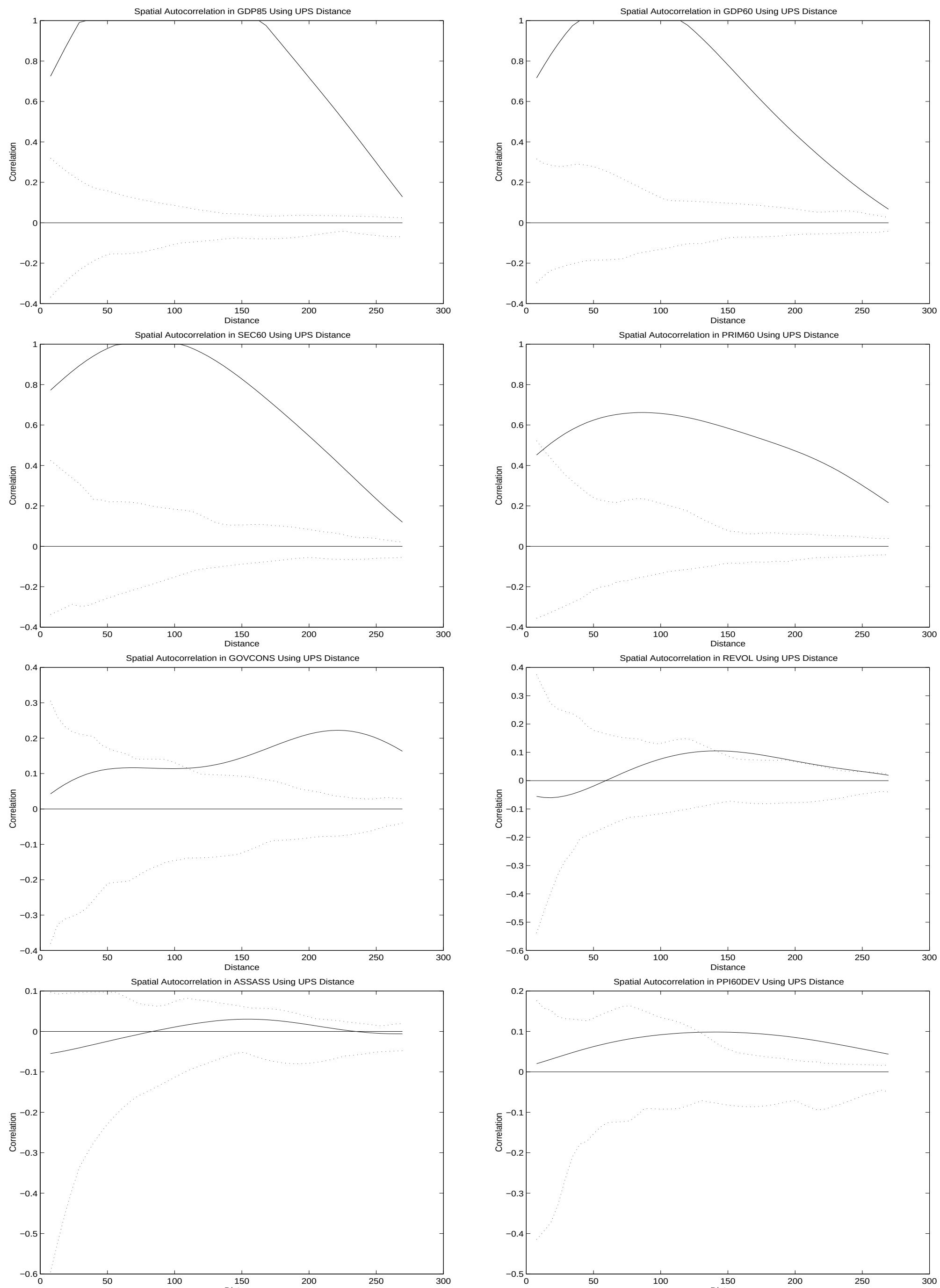

FIGURE 6. Spatial Autocorrelation of Various Variables, using UPS Distance. The dotted regions in each figure are bootstrapped 95 per cent acceptance regions for the null hypothesis of independence.

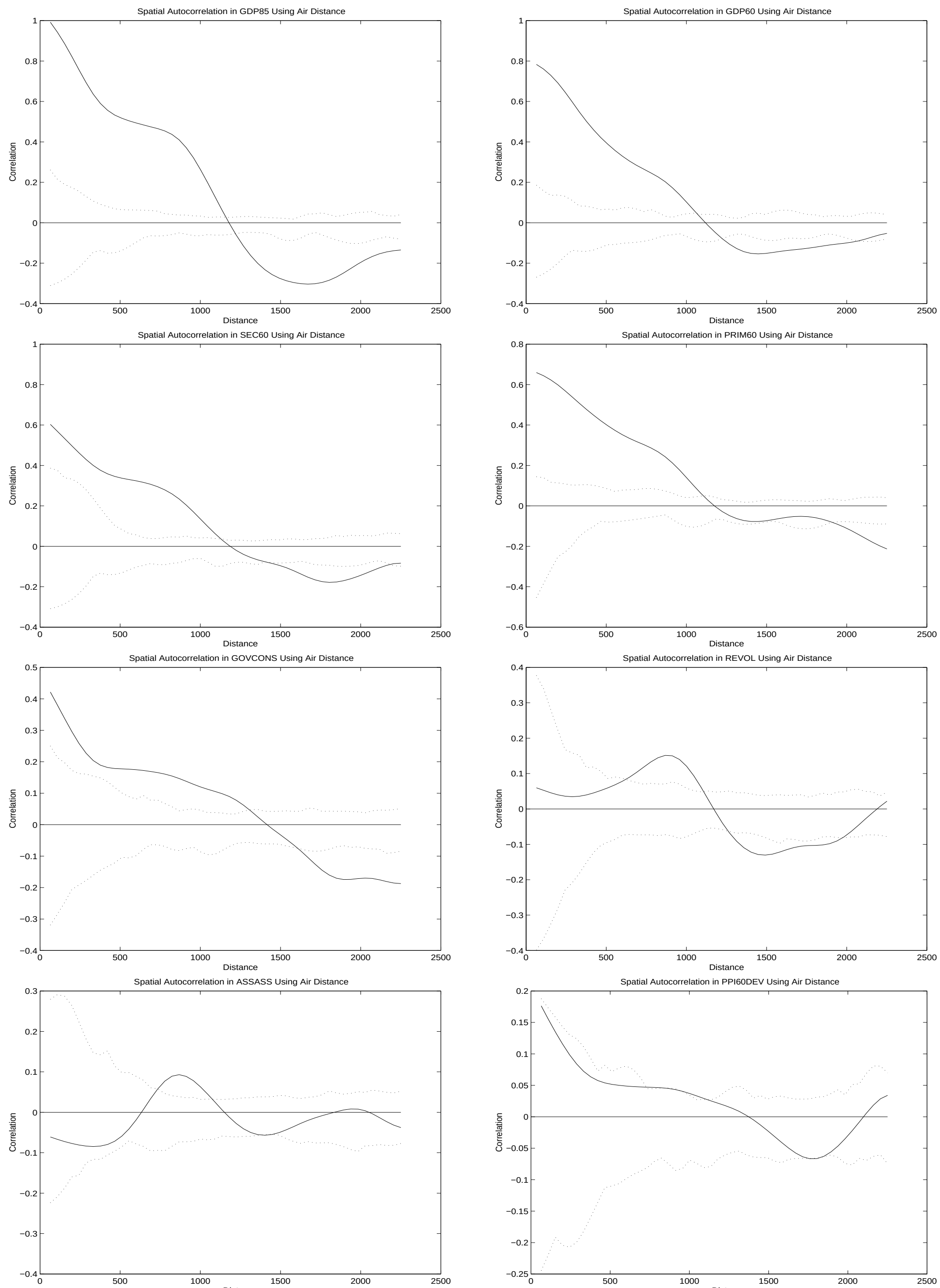

Figure 7. Spatial Autocorrelation of Various Variables, using Airfare Distance. The dotted regions in each figure are bootstrapped 95 per cent acceptance regions for the null hypothesis of independence.

above the ACF for 1960, though the shape of the ACF changes very little. Some other variables are notable for their general lack of spatial dependence; these include the social unrest variable ASSASS, as well as PPI60DEV.

4.2. **Growth Regressions.** In this section we investigate whether the spatial correlation in growth rates can be rationalized by the fact that many of the economic variables commonly thought of as determinants of growth are spatially correlated. We estimate a growth regression and test whether there remains spatial correlation in the residuals. We use OLS to estimate the parameters of specification (14) of Barro (1991); using the variables described above, along with a constant and dummy variables for Latin America and sub-Saharan Africa. OLS is a just identified GMM estimator so allowing for correlation as a function of economic distance will only impact inference, not point estimates. Thus, our point estimates replicate one of the cross-country growth regressions of Barro (1991);[14] however, our estimates of the standard errors will differ from Barro's, since we allow for spatial dependence. This focus on inference permits us to keep comparisons simple; the impact of allowing for dependence will likely be greater in over-identified systems, since in this latter case estimated asymptotic covariance matrices will determine optimal weighting matrices and hence point estimates and tests of over-identifying restrictions as well as standard errors.

We regressed the growth rate of GDP from 1960 to 1985 on a constant, GDP in 1960 (GDP60), secondary and primary school enrollment rates (SEC60 and PRIM60), the share of real government consumption expenditures to real GDP for 1970–1985 (GOVCONS), revolutions from 1960–1985 (REV), assassinations per million (ASSASS), and the absolute deviation of 1960 PPP investment deflator from its sample mean (PPI60DEV), and dummies for Africa(AFRICA) and Latin America (LAAMER). This is exactly specification (14) of Barro (1991). Table 2 presents OLS point estimates and standard errors of these estimates calculated in five different ways. First, the usual OLS standard errors are labelled IID SE, then heteroskedasticity consistent standard errors are labelled White SE (White 1980) (these are the standard errors reported in Barro (1991)). The next three columns of standard errors are calculated using our three different measures of economic distance to place some structure on the cross-sectional dependence of

[14]While we nearly replicate Barro's results, the point estimates we report differ slightly from those reported in Barro (1991) because our sample is slightly different because we lack distance data of one sort or another for Liberia, Mauritius, Rwanda, Swaziland, Myanamar, Iran and Malta.

Geographic Distance						
Variable	Point Est.	IID SE	White SE	1300 mi.	2100 mi.	2900 mi.
Constant	0.0338	0.0063	0.0069	0.0068	0.0050*	0.0019*
GDP60	-0.0067	0.0011	0.0009	0.0009	0.0008	0.0009
SEC60	0.0120	0.0106	0.0077	0.0046*	0.0057*	0.0046*
PRIM60	0.0273	0.0060	0.0060	0.0059	0.0033*	0.0052
GOVCONS	-0.0971	0.0260	0.0266	0.0315	0.0302	0.0264
REVOL	-0.0208	0.0085	0.0081	0.0086	0.0104*	0.0108*
ASSASS	-0.0025	0.0029	0.0018	0.0016	0.0019	0.0013*
PPI60DEV	-0.0136	0.0051	0.0047	0.0041	0.0038*	0.0032*
AFRICA	-0.0110	0.0038	0.0040	0.0032*	0.0034	0.0028*
LAAMER	-0.0139	0.0033	0.0031	0.0011*	0.0019*	0.0016*
UPS Distance						
Variable	Point Est.	IID SE	White SE	$275	$325	$375
Constant	0.0338	0.0063	0.0069	0.0072	0.0055*	0.0063
GDP60	-0.0067	0.0011	0.0009	0.0011	0.0010	0.0007*
SEC60	0.0120	0.0106	0.0077	0.0076	0.0060*	0.0053*
PRIM60	0.0273	0.0060	0.0060	0.0056	0.0062	0.0062
GOVCONS	-0.0971	0.0260	0.0266	0.0308	0.0296	0.0317
REVOL	-0.0208	0.0085	0.0081	0.0077	0.0079	0.0029*
ASSASS	-0.0025	0.0029	0.0018	0.0022	0.0020	0.0021
PPI60DEV	-0.0136	0.0051	0.0047	0.0047	0.0049	0.0051
AFRICA	-0.0110	0.0038	0.0040	0.0038	0.0028*	0.0027*
LAAMER	-0.0139	0.0033	0.0031	0.0033	0.0029	0.0026*
Air Fare Distance						
Variable	Point Est.	IID SE	White SE	$800	$1000	$1200
Constant	0.0349	0.0063	0.0070	0.0081	0.0070	0.0018*
GDP60	-0.0067	0.0011	0.0009	0.0011	0.0009	0.0009
SEC60	0.0145	0.0107	0.0074	0.0083	0.0074	0.0065
PRIM60	0.0256	0.0063	0.0063	0.0059	0.0052*	0.0035*
GOVCONS	-0.1058	0.0264	0.0273	0.0350*	0.0300	0.0245
REVOL	-0.0204	0.0085	0.0081	0.0069	0.0076	0.0073
ASSASS	-0.0025	0.0029	0.0017	0.0020	0.0019	0.0013*
PPI60DEV	-0.0133	0.0051	0.0048	0.0049	0.0047	0.0047
AFRICA	-0.0108	0.0040	0.0044	0.0043	0.0041	0.0045
LAAMER	-0.0132	0.0033	0.0031	0.0025*	0.0017*	0.0022*

TABLE 2. Estimates and Standard Errors for a Growth Regression. The bandwidths indicated in the column headings correspond to roughly 9%, 17%, and 24% of the total country pairs for geographic distance; the bandwidths for UPS distances correspond to 10%,20%, and 33% of total country pairs; and airfare distances correspond to 9%, 15%, and 23% of all country pairs. Asterisks indicate a difference of at least 20 per cent relative to the White standard errors.

the OLS residuals. For these estimated standard errors, we use the estimator in Section 3.2 for a range of bandwidths in order to give some sense of the robustness of these estimates.

The estimates in Table 2 provide evidence that allowing for economic distance can be important for conducting inference. Overall, the magnitude of the difference between spatial standard errors is at least as great as the difference between IID standard errors and White standard errors; asterisks indicate a difference of at least 20 per cent. The outcome of t-tests changes for several variables at conventional levels of significance, most importantly SEC60.

Table 2 illustrates another important point: that spatial dependence does not imply that standard errors will rise. Indeed, many of the spatial standard error estimates are smaller than their IID or White counterparts. It is important to remember that asymptotic variances may be smaller with spatially dependent data, just as asymptotic variances can be lower for dependent time series averages than independent series averages. This is a direct consequence of negative spatial correlation in residuals, shown in Figure 8. This case is analogous to the average of a time series with negative serial correlation having smaller asymptotic variance than a serially independent series with the same stationary (marginal) distribution.

Figure 8 displays computed spatial autocorrelation functions of residuals using each of the three distance measures discussed in Section 2, along with the weighting function described in Section 4.1. Bandwidths are 300 miles for geographic distance, \$50 for UPS distances, and \$150 for airfares.

The positive spatial autocorrelation evident in growth rates doesn't seem to be a feature of the residuals from our estimating equation. While there's weak evidence of positive spatial dependence when we use UPS distance, there's actually surprising evidence of at least some negative spatial dependence when using geographic distance.

Despite some evidence of spatial dependence in the OLS residuals, probably the most striking feature of Figure 8 relative to the spatial autocorrelations shown for growth rates in Figure 4 is how nearly uncorrelated the residuals are. The OLS regression described above is surprisingly successful in explaining the spatial relationship between growth rates.

4.3. **Spillovers.** We investigate the potential for spillovers to explain cross-country growth patterns by testing whether neighbors' variables have additional power to predict a country's growth rate beyond that country's own characteristics. We simply look at whether residuals

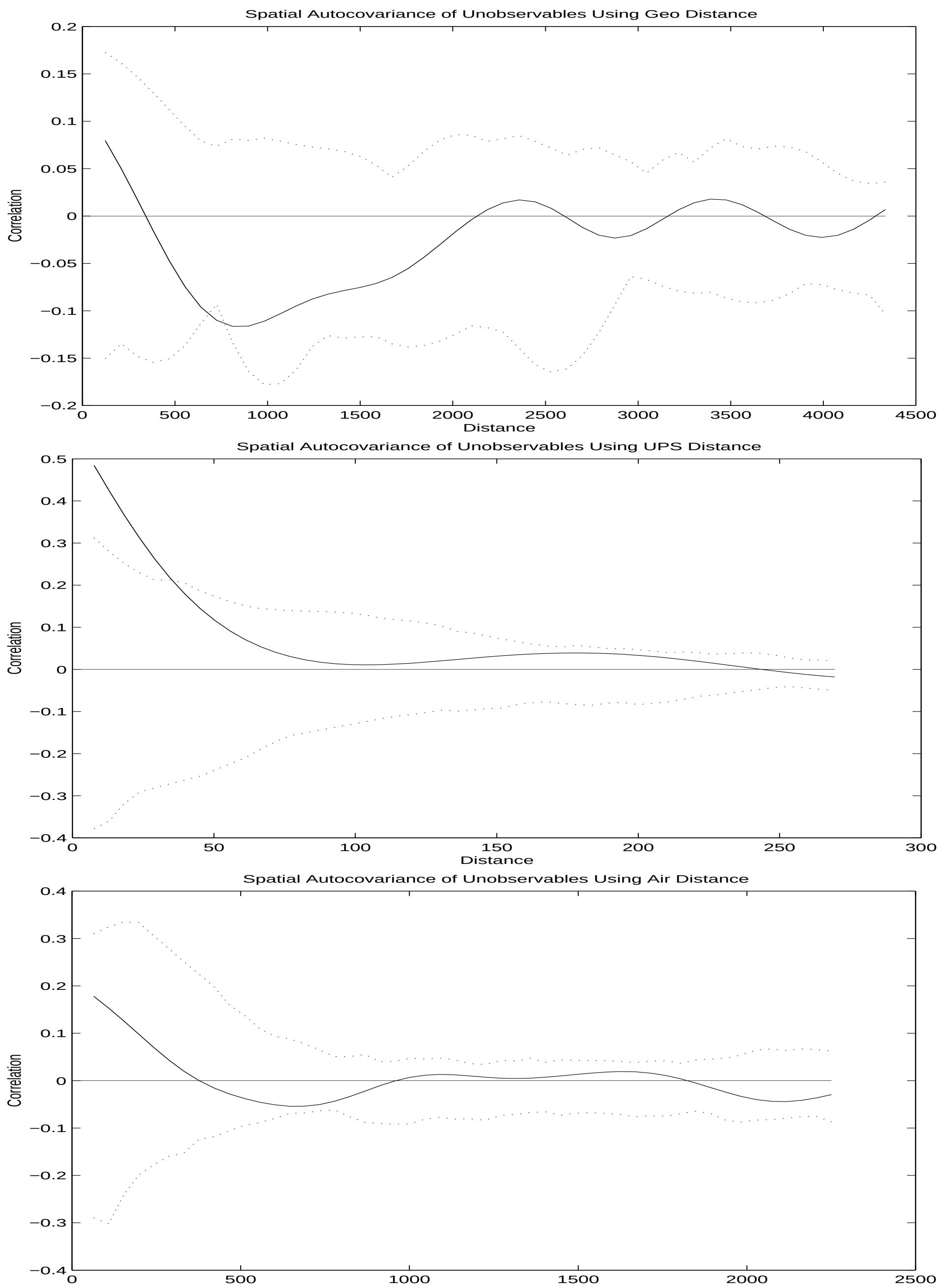

FIGURE 8. Spatial autocorrelation of residuals. The first panel uses geographic distance, the second UPS distance, and the third air fare distance. The dotted regions in each figure are bootstrapped 95 per cent acceptance regions for the null hypothesis of spatial independence.

from the growth regressions estimated above are correlated with other countries' variables as a function of economic distance. If neighbors' variables are correlated with the residuals from a growth regression, then this is evidence that those variables would have power in the growth regression. This method enables us to test for a general presence of spillover-type effects without the need to specify a particular 'spillover regressor' in a regression.

For each country we can write its growth rate as the sum of its predicted value $\hat{y}_{s_i}$ and a residual $\hat{u}_{s_i}$, or

$$y_{s_i} = \hat{y}_{s_i} + \hat{u}_{s_i}.$$

Variables that are correlated with the prediction errors $\hat{u}_{s_i}$ would improve predictions of y_{s_i}. Thus we can investigate the potential for any of neighboring countries' variables to be useful in predicting y_{s_i} by testing whether there is spatial covariance between the variable and $\hat{u}_{s_i}$. Of course, this test has the drawback of just looking at one particular variable or combination of variables at a time so it may have low power to detect some spillovers. We estimated spatial covariances between our residuals $\hat{u}_{s_i}$ and all the continuous regressors in the regression specified above. Figure 9 contains our estimates of the spatial correlation between PRIM60 and GDP60 as a function of our three economic distances along with acceptance regions for the null hypothesis of independence. The results for these variables indicate that there is signficant predictive power in neighbors' values of these two variables when the airfare or UPS metrics are used. However, when the metric of physical distance is used there is no evidence of neighbors' variables having predictive power. We obtain almost identical qualitative results for the variables SEC60, GOVCONS, and PPI60DEV. Evidence of spillovers is strongest when we use UPS and airfare distances, rather than geographic distance—the least reflective of the economic costs.

While the statistically significant covariances described above offer evidence of spillovers being non-zero, there is still the question of whether they are generally large or small in magnitude. In order to address this question we look at the spatial covariance between our regression residuals and a particular linear combination of our determinants-of-growth variables. We choose to use the linear combination corresponding to the predicted values for each country, $\hat{y}_{s_i}$. This enables us to portray the 'spillover covariance' relative to the covariance of the growth rates, predicted values, and residuals. We can write the covariance of y_{s_i} and y_{s_j} as

$$\operatorname{cov}(y_{s_i}, y_{s_j}) = \operatorname{cov}(\hat{y}_{s_i}, \hat{y}_{s_j}) + \operatorname{cov}(\hat{u}_{s_i}, \hat{u}_{s_j}) + 2\operatorname{cov}(\hat{y}_{s_i}, \hat{u}_{s_j}).$$

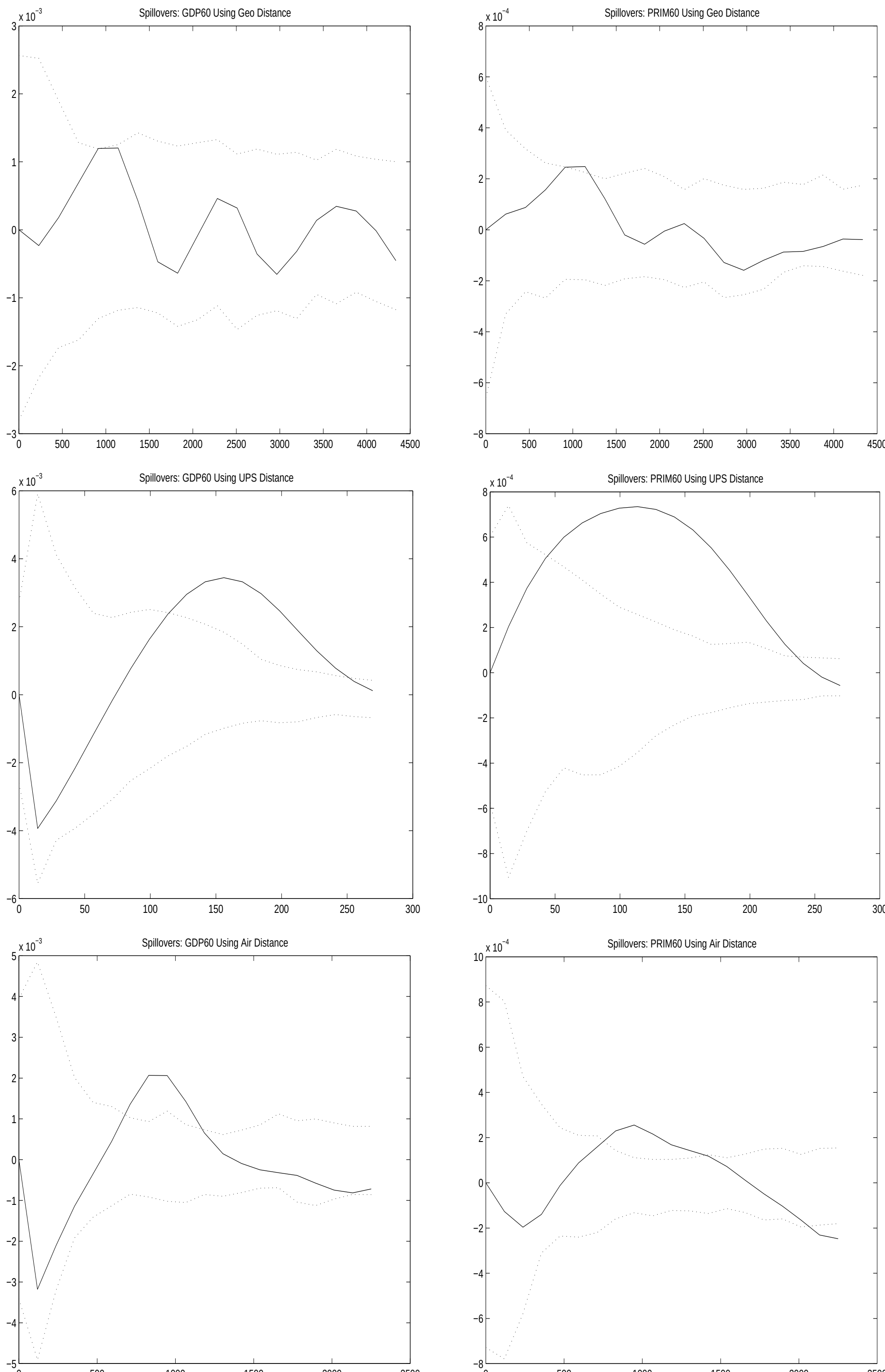

FIGURE 9. A selection of spillovers. The first row uses geographic distance; the second UPS distance; and the third airfares. Dotted lines in each figure form bootstrapped 95 per cent acceptance regions for the null hypothesis of spatial independence.

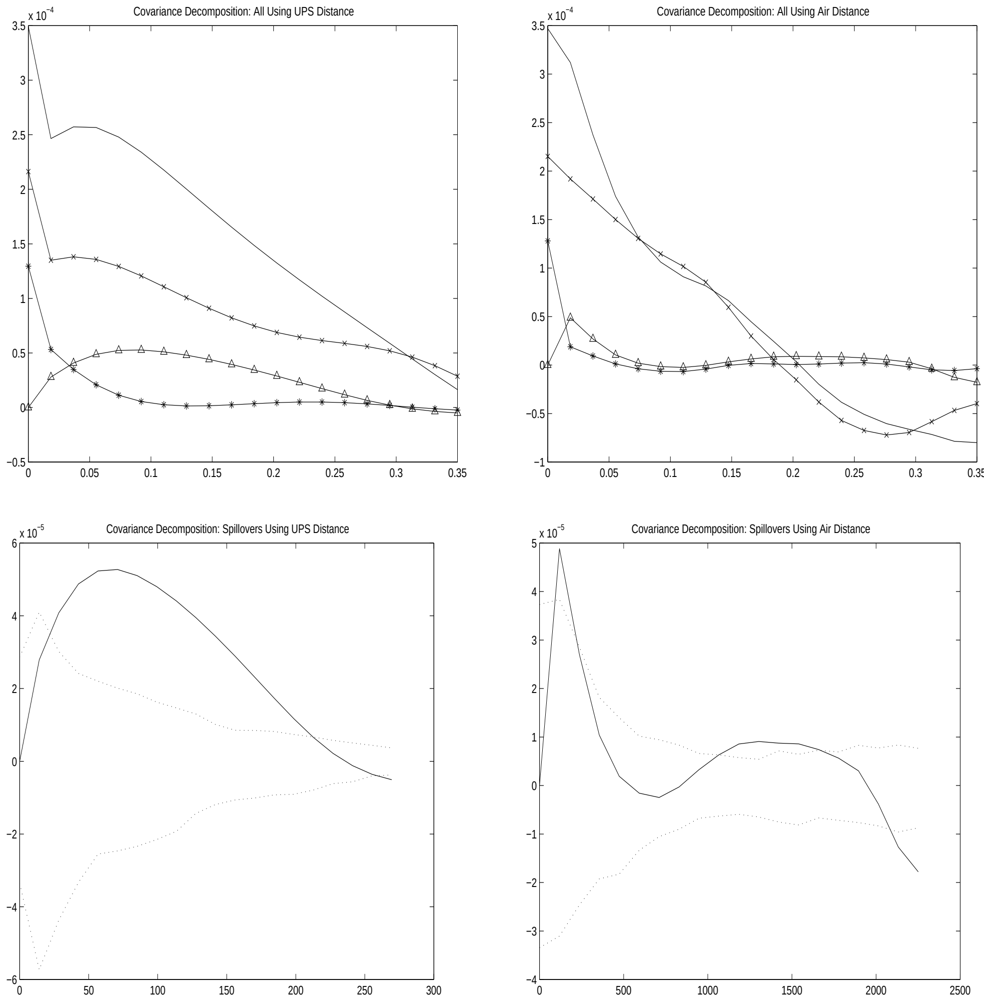

FIGURE 10. Covariance Decomposition. Dotted lines in each figure form bootstrapped 95 per cent acceptance regions for the null hypothesis of spatial independence.

Thus, decomposing the spatial covariance of y into covariances of three terms: observables (predicted values), unobservables (residuals), and spillovers. Figure 10 contains a representation of these components for the airfare and UPS metrics. These figures demonstrate that the spillovers are not only statistically significant, but are also appreciable in magnitude relative to the spatial correlation in growth rates and predicted growth rates.

5. Conclusion

In this paper, we have described several ways for economic distances to inform the analysis of cross-country data. The two basic methods are estimating spatial correlations as functions of economic distances between countries and allowing for cross-sectional dependence as a function of economic distance in GMM estimation. These methods will be useful in any application where some (perhaps imperfect) measure of economic distance between countries is available. We illustrated these methods in an investigation of cross-country patterns in growth rates using three measures of economic distance between countries—geographic, UPS, and airfare.

We used the structure provided by economic distance measurements to estimate the spatial autocorrelation of variables of interest. We found significant spatial autocorrelation of GDP growth rates and a variety of other economic variables commonly regarded as determinants of growth.

In order to investigate whether the cross-country correlation patterns in growth rates could be explained by correlation in observable explanatory variables we estimated a growth regression suggested by Barro (1991). We used our spatial model to allow for dependence across countries in drawing inferences from this cross-country regression. Our inferences about variables' impact on growth were different when we allowed for cross sectional dependence as a function of our candidate economic distances. Inferences about the impact of secondary education on growth were perhaps most sensitive as significance tests reversed their conclusions. We found little evidence of spatial correlation in the residuals from this regression. This leads us to the conclusion that cross-country correlation in these explanatory variables can indeed account for most of the cross-country correlation in growth rates.

Finally, we used our economic distance measurements to address the question of whether spillovers are important determinants of growth. For two of our candidate economic distance metrics (the UPS and airfare metrics) a country's neighbors' characteristics offered additional explanatory power in predicting its growth rate, beyond that of its own characteristics. We did not find evidence of such spillover effects when neighbors were defined using geographic distance.

References

Barro, R. J. (1991). Economic growth in a cross section of countries. *Quarterly Journal of Economics 106*, 407–444.

Conley, T. G. (1996). Econometric modelling of cross sectional dependence. Manuscript, Northwestern University.

DeLong, J. B. and L. H. Summers (1991). Equipment investment and economic growth. *Quarterly Journal of Economics 106*(2), 445–502.

Elliott, G. (1993). Spatial correlation and cross country regressions. Manuscript.

Hall, P., N. I. Fisher, and B. Hoffman (1992). On the non-parametric estimation of covariance functions. Manuscript, Australian National University.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50*, 1029–1054.

Levine, R. and D. Renelt (1992). A sensitivity analysis of cross-country growth regressions. *American Economic Review 82*(4), 942–963.

Quah, D. (1992). International patterns of growth: I. persistence in cross-country disparities. Manuscript.

Quah, D. (1993). Galton's fallacy and tests of the convergence hypothesis. Manuscript.

Summers, R. and A. Heston (1991). The Penn World Tables (Mark 5): An expanded set of international comparisons. *Quarterly Journal of Economics 106*(2), 327–368.

White, H. (1980). A heteroskedastic-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica 48*, 817–38.

Northwestern University

University of California, Berkeley