CHRIS EDMOND
Reserve Bank of Australia
Sydney, Australia and
University of California, Los Angeles
Los Angeles, California

*Some Panel Cointegration Models of International R&D Spillovers**

Coe and Helpman (1995) estimated a relationship between TFP and levels of domestic and foreign R&D capital, but couldn't provide compelling evidence of the panel cointegration needed to support their estimation strategy. This paper uses Pedroni's (1997, 1998) tests for panel cointegration in both the Coe-Helpman setup and in a framework with more cross-section heterogeneity. Criticisms of the Coe-Helpman approach are also tested. Coe and Helpman's models exhibit cointegration, but when more heterogeneity is allowed, coefficient estimates appear less robust. The elasticity of productivity with respect to foreign R&D is unstable across alternative estimation strategies.

1. Introduction

In a recent study of 21 OECD countries and Israel over 1971–1990, Coe and Helpman (1995) used a simple single equation framework to estimate the relationship between a country's level of total factor productivity and levels of domestic and foreign R&D capital stocks. Briefly, Coe and Helpman found evidence to support the idea that foreign R&D has strongly beneficial effects for domestic productivity and that the extent of the benefit increases with the degree of openness of the domestic economy to international trade. Since Coe and Helpman analyzed the reduced form of a long-run equilibrium relationship, it was natural for them to try to find a cointegration representation. As part of their analysis, Coe and Helpman presented the results of Levin and Lin (1992, 1993) panel unit root tests. These tests were employed both to test the stationarity of the underlying panels and as a residual-based test for cointegration. Coe and Helpman also used a single equation error correction test for cointegration (that is, a test of the significance of lagged residuals in the error correction representation).

*I would like to thank Peter Pedroni for permission to use his code. Phil Bodman and three anonymous referees provided valuable comments on and criticisms of earlier versions of this paper. All remaining errors and omissions are my own, however. The opinions expressed in this paper are not necessarily those of the Reserve Bank of Australia. My code and the results from some preliminary analysis mentioned in the text are available on request. E-mail: ⟨cedmond@ucla.edu⟩

0164-0704/2001/$1.50

However, the results of Coe and Helpman's time series analysis were inconclusive. One of their models—though not their preferred model—exhibited cointegration, but only according to the version of the Levin-Lin test least likely to be applicable to their panel data.[1] The error correction test for cointegration indicated that all of their models were cointegrated.[2] Eventually, Coe and Helpman (1995, 870) concluded their analysis by remarking:

> Given these mixed results, and given that the econometrics of pooled cointegration are not yet fully worked out, we place more emphasis on the a priori plausibility of the estimated parameters than on the tests for cointegration.

The econometrics of pooled and panel cointegration are still not fully worked out. However, recent advances in econometric theory (that were not available at the time of Coe and Helpman's research) allow the cointegration issue to be pursued further. The first half of this paper is broadly concerned with these pre-testing issues and employs the panel cointegration procedures developed by Pedroni (1997, 1998). Second, this paper presents a set of alternative results from an estimation method that allows considerably more cross section heterogeneity than was allowed by Coe and Helpman. Alongside the treatment of these main concerns, parallel consideration is given to the time series properties of the alternative definitions of the foreign R&D capital stock used by Keller (1998) and Lichtenberg and von Pottelsberghe de la Potterie (1998; LP hereafter).[3]

The results reported here suggest that Coe and Helpman's intuition was correct. Their models are cointegrated, at least according to their preferred pooled data methods. However, when group mean methods are used (which allow for more cross section heterogeneity), the results of Coe and Helpman's approach look less robust. The elasticity of productivity with respect to foreign R&D is not stable across alternative estimation strategies. It is also found that LP's recommended reconstruction of Coe and Helpman's preferred equation cannot be part of a cointegrating relationship since LP's alternative measure of a country's foreign R&D capital stock is (trend) stationary.

[1]That is, the Levin and Lin (1992) test; a test that requires homogeneity of both the autoregressive root under the alternative hypothesis and of autocorrelation structures across cross section members.

[2]Although Coe and Helpman (1995, 869) reported the "wrong" critical values for their error correction tests—they did not take account of the way fixed effects dummies alter the distribution of the test statistic (see Kremers, Ericsson and Dolado 1992)—the same conclusions are obtained even when the "correct" (i.e., conservative Dickey-Fuller) critical values are used.

[3]Neither Keller nor LP reported the results of any time series analysis of their alternative models and/or measures of foreign R&D capital stocks.

A review of the literature's basic approach and results is given in Section 2. In Section 3 this paper revisits Coe and Helpman's analysis of the stationarity properties of their data and presents some alternative group mean panel unit root tests. Section 3 also considers some of the alternative foreign R&D capital stock measures discussed by Keller (1998) and LP. Section 4 presents the results of Pedroni's (1997, 1998) panel cointegration tests using both the pooled data and group mean approaches, while Section 5 concludes and offers some brief suggestions for further research.

2. A Brief Review

In essence, Coe and Helpman's approach to international R&D spillovers is to estimate a single equation of the form:

$$\ln F_{it} = \alpha_i + \alpha^d \ln S^d_{it} + \alpha^{G7}\, G7 \ln S^d_{it} + \alpha^f(M_{it}/Y_{it}) \ln S^f_{it} + \epsilon_{it}\,, \quad (1)$$

where F_{it} is total factor productivity, S^d_{it} is a stock of R&D capital found by accumulating domestic R&D expenditures, $G7$ is a dummy variable for those seven countries, (M_{it}/Y_{it}) is a country's ratio of total imports to output, S^f_{it} is a country's stock of foreign R&D capital and ϵ_{it} is an idiosyncratic disturbance.[4] There are 21 OECD countries and Israel over 1971–1990 at the annual frequency (i.e., the panels are 22×20). Coe and Helpman's results were that the elasticity of productivity with respect to domestic R&D, $\alpha^d + \alpha^{G7}G7$, was about 0.23 for the G7 countries and about 0.08 for the remaining 14 countries in their sample. The elasticity of productivity with respect to foreign R&D, $\alpha^f\,(M_{it}/Y_{it})$, ranges from about 0.02 to 0.08 for the G7 countries and from 0.04 to 0.26 for the others. For both G7 and non-G7 countries, the elasticity with respect to foreign R&D was found to be generally rising over time.[5]

Much of the contention surrounding Coe and Helpman's results lies in the specification of S^f_{it}. Coe and Helpman's original approach was to compute S^f_{it} by weighting the domestic R&D stocks of country i's partners by bilateral import shares, m_{ijt}:

$$S^f_{it} = \sum_{j \neq i} m_{ijt} S^d_{jt}\,.$$

The foreign R&D stock then represents a hypothesis about the effect of a

[4]Coe and Helpman first consider versions of Equation (1) without the G7 dummy and without the use of the aggregate import share, but (1) represents their preferred specification. See Coe and Helpman (1995) and Keller (1998) for discussion of the theoretical origins of this approach.

[5]Norway, Spain and Switzerland were the exceptions to this trend.

country's *composition* of trading partners on its productivity—other things equal, a country that imports mostly from high R&D countries should benefit more than a country that imports mostly from low R&D countries. The additional use of the ratio (M_{it}/Y_{it}) then represents a hypothesis about a country's *intensity* of trade; even if two countries have the same composition of trading partners, a country with a higher aggregate openness to imports should benefit more than a country with a lower openness to imports.

Keller (1998) shows that regressions using counter-factual foreign R&D stocks (computed using certain randomized bilateral trade shares) can produce equally large or even larger R&D spillovers than were obtained by Coe and Helpman's method.[6] Keller (1998) also shows that using a simple, un-weighted, sum of domestic R&D stocks produces even higher estimated R&D spillovers. Both of these results cast doubt on the composition of trade hypothesis. Hence, this paper also employs the simple-sum versions of Keller's regressions, as they are easier to work with than the counter-factual foreign R&D stocks.[7]

LP show that Coe and Helpman's original version of (1) was mis-specified. Coe and Helpman measure S^f_{it} as index numbers, so after taking logarithms the base of the index number times the import share, $(M_{it}/Y_{it})\ln S^f_{i,1985}$, is time-varying and (unlike the bases of the index numbers in S^d_{it}) cannot be incorporated into the fixed effects dummies, α_i.[8] As noted by LP (1998: 1486–90) and Coe and Hoffmaister (1999: 11), a reformulated version of Equation (1) needs to add the import share, (M_{it}/Y_{it}), as an additional regressor to avoid imposing cross restrictions on the elasticities of productivity with respect to the import share and with respect to foreign R&D capital. Instead of the specification in (1), we have:

$$\ln F_{it} = \alpha_i + \alpha^d \ln S^d_{it} + \alpha^{G7}\, G7 \ln S^d_{it} + \alpha^f (M_{it}/Y_{it}) \ln S^f_{it} + \alpha^m (M_{it}/Y_{it}) + \epsilon_{it} \,.$$

[6]By contrast, Keller (1997) shows that certain other randomized trade shares lead to no significant coefficient for foreign R&D, and Coe and Hoffmaister (1999) demonstrate that this is in fact true for most randomizations of trade patterns in the Coe-Helpman (1995) data set. Nevertheless, Keller's (1998) regressions using counter-factual trade patterns still show that the Coe-Helpman findings did not rely on using the actual import patterns.

[7]Coe and Helpman (1995, 884) report that they experimented with a simple sum definition of foreign R&D capital stocks, but they do not present the results from these regressions because they prefer the trade weighted definition on theoretical grounds. Coe, Helpman and Hoffmaister (1997) also present results using a simple sum definition of R&D.

[8]Coe and Hoffmaister (1999, 12) correct for this mis-specification and show that making this correction has little effect on Coe and Helpman's original results. For example, in the regression corresponding to Equation (1), the elasticity of productivity with respect to domestic R&D goes from about 0.234 to 0.237 for the G7 and from 0.078 to 0.082 for the non-G7.

LP (1998, 1484–7) also show that the intensity weighted foreign R&D stock is sensitive to the level of data aggregation and will necessarily give misleading results in the presence of mergers between countries. They propose the alternative measure:

$$S_{it}^{f-LP} = \sum_{j \neq i} \frac{m_{ijt} S_{jt}^{d}}{Y_{jt}} \,. \tag{2}$$

LP argue that estimation using a version of Equation (1) where $\ln S_{it}^{f-LP}$ replaces $(M_{it}/Y_{it}) \ln S_{it}^{f}$ provides a better measure of the elasticity of productivity with respect to foreign R&D capital.

All of the procedures discussed in this section estimate a panel regression like (1) by pooling data into one consolidated vector per variable and use fixed effects dummies to control for the mean differences between the sub-samples that represent each country. This approach imposes homogeneity of slope coefficients and error variances.[9] Although pooling is a common way to deal with panel data when stationarity issues are not of any concern, it is not necessarily innocuous with non-stationary time series. (In particular, in a dynamic panel data setting where lagged dependent variables enter as regressors, pooling does not produce consistent estimates unless the true slope coefficients are identical over the cross section members [see Pesaran and Smith 1995, 79–81].)

Cointegration in a panel setting raises a related set of issues. Coe and Helpman use Levin-Lin (1992, 1993) panel unit root tests on the obtained residuals from specifications like Equation (1). For ordinary time series data—and given a slight adjustment of critical values to reflect the estimated nature of the residuals—this approach constitutes a valid test of the null hypothesis of no cointegration. However, in the panel data setting, Pedroni (1997) has shown that this application of a "raw" panel unit root test is only valid if slope coefficients are truly the same across members *and* the dynamics of the error processes are the same across members. If these conditions are not met, then the true critical values of the panel cointegration test are far to the left of the critical values that are valid for the dynamic homogeneity case.[10] For example, a panel unit root test with a 10% critical statistic of -1.28 equates to a panel cointegration critical statistic of -8.71 (Pedroni

[9]As in a SUR model, however, between-individual differences in error variances could be handled with a simple feasible GLS approach (e.g., a block diagonal structure for the variance-covariance matrix of the errors could be imposed).

[10]If the dynamic homogeneity requirements *are* met, then Pedroni (1997) shows that it does not matter whether the residuals used are estimated or known.

1997, 3).[11] This paper exploits recent results in both homogeneous and heterogeneous panel cointegration theory to present estimates both in the original Coe-Helpman pooled data framework and in a group mean framework that allows for more heterogeneity across countries.

3. Panel Unit Root Preliminaries

In general, panel unit root tests have substantially more power against near-integrated alternatives than the usual time series tests. The Levin-Lin (1992, 1993) panel unit root tests used by Coe and Helpman correspond to a pooled approach to panel unit root testing where the null hypothesis is that each member of the panel has a unit autoregressive root and the alternative is that the members share a common root, $\rho_i = \rho$, less than one. In Levin and Lin (1992), autocorrelation structures are constrained to be the same across all cross section members, while Levin-Lin (1993) permits autocorrelation structures to vary across members. Since Coe and Helpman present the results of both kinds of Levin-Lin test, this paper does not replicate those tests. However, one of the reasons for being interested in a panel unit root test—apart from increased power—is that such a test might provide a kind of summary of the time series properties of the variable of interest

[11]Pedroni's results are based on sequential limit theory (i.e., fixing N, letting $T \to \infty$, then subsequently letting $N \to \infty$). See Phillips and Moon (1999b) for analysis of the relationship between this kind of limit theory and joint limit theory where both $N, T \to \infty$ (possibly with a rate condition like $T/N \to 0$ governing the expansion). Phillips and Moon also investigate conditions for which sequential and joint limits are equivalent. The diagonal path limit theory used by Quah (1994), Levin and Lin (1992, 1993), and Im, Pesaran and Shin (1997) in deriving their various panel unit root tests is a special case of joint limit theory. See also Kao (1999) for an alternative approach to spurious regression and cointegration in panel data. The pooled fully-modified estimators proposed by Phillips and Moon (1999a, 1999b) trade off more robust asymptotic theory for more restrictive assumptions (usually requirements for certain higher order moments to exist) than are needed for sequential limits. Since fully modified estimators (e.g., Phillips and Hansen 1990) are usually used to purge endogeneity from the regressors and to overcome the finite sample bias associated with the mis-specified dynamics of an Engle-Granger approach to cointegration, Phillips and Moon's results might initially seem to provide a suitable alternative approach to studying the properties of these models. However, a fully modified approach also requires that the *regressors* themselves are not cointegrated, yet in this literature the foreign R&D capital stocks are *by construction* a linear combination of the domestic R&D capital stocks and so it seems a priori likely that the regressors may in fact be cointegrated. Also, since this paper never performs any formal statistical inference (e.g., hypothesis tests) on obtained coefficients, there is no technical requirement for a fully modified approach. Hence, for the purposes of this paper, attention has only been given to the results from Pedroni's procedures. Nonetheless, the section below titled "Discussion of Estimation Results" indicates how the use of Phillips and Moon's fully modified approach may help in interpreting some of the results obtained in this paper and may therefore be a useful approach for future work on the time series properties of R&D spillovers models.

across individual panel members. The Levin-Lin tests cannot provide this kind of summary information because of the way the alternative hypothesis has to be specified, but the test developed by Im, Pesaran and Shin can. Table 1 shows Im, Pesaran and Shin (1997; IPS hereafter) group mean panel unit root tests (or "t-bar" tests), which allow each member of the cross section to have a different autoregressive root *and* different autocorrelation structures under the alternative hypothesis.[12]

The results reported in Table 1 all use the data from Coe and Helpman's Appendix A. Since Coe and Helpman only report a bilateral trade matrix for one year, 1990, this paper uses that static matrix for the calculations that involve bilateral trade weighting.[13] The mean t statistic on the autoregressive coefficient is $\bar{t}$, $\tilde{p}$ is the mean minimum number of differences needed to purge autocorrelation from the ADF regressions that are run for each country, while the mean and variance adjustment terms are the averages of the individual adjustment terms reported in IPS.[14] From here on, this paper uses the notation $\ln S^f_{it}$ for the logarithm of the raw level of S^f_{it}, not the original index number—which is denoted S^{f-CH}_{it}. The variable S^s_{it} is the simple sum measure of the foreign R&D capital stock, as used by Keller. IPS show that the group mean test statistic has a standard normal distribution, and as usual, significantly negative test statistics indicate rejection of the unit root null hypothesis. Table 1 indicates that the null can only be rejected for $\ln S^{f-LP}_{it}$.[15,16] Since the null hypothesis that the $\ln S^{f-LP}_{it}$ panel has a unit root is strongly rejected, regressions using $\ln S^{f-LP}_{it}$ will not meet one of the necessary conditions for cointegration. This is important because one of the attractions of a cointegrating relationship is that it allows a re-

[12]See Quah (1994), Levin and Lin (1992, 1993), Im, Pesaran and Shin (1997) for details of these and related tests and for the development of the relevant asymptotic theory. A brief description of the mechanics is given in the Appendix.

[13]Keller (1998) reports from his simulation study that it doesn't seem to matter whether time-varying bilateral trade shares are used or not.

[14]For each individual in the cross section, the lag order p_i was selected by finding the minimum number of lags required to purge autocorrelation from the ADF regression. The most parsimonious representation was selected using the SBC criterion. The results of this preliminary ARIMA analysis, including details of the individual lag orders selected, are available on request.

[15]One explanation for the trend stationarity of $\ln S^{f-LP}_{it}$ might be guessed at from Equation (2). LP's measure is constructed by dividing S^d_{jt} by Y_{jt}. To the extent that cross section national outputs each contain substantial permanent components, this is (loosely speaking) akin to stochastically detrending S^d_{jt}. If so, we have the product of two stationary variables, m_{ijt} and (S^d_{jt}/Y_{jt}), which will itself be stationary.

[16]Table 1 in an earlier version of this paper mistakenly reported that $\ln S^f_{it}$ was also trend stationary, despite the fact that S^f_{it} is a scalar multiple of S^{f-CH}_{it}, such that $\ln S^f_{it}$ must be nonstationary if $\ln S^{f-CH}_{it}$ is. This embarrassing mistake was fortunately spotted by an anonymous referee.

TABLE 1. *Group Mean Panel Unit Root Tests (Im, Pesaran and Shin 1997 tests)*

Variable	$\bar{t}$ (a)	$\bar{p}$ (b)	Mean Adjustment (c)	Variance Adjustment (d)	Group Mean Statistic (e)	Decision (f)
$ln\ F_{it}$	−2.3221	1.5000	−2.1178	0.8710	−1.0267	$I(1)$
$ln\ S_{it}^{d}$	−1.8339	1.5455	−2.0985	0.8770	1.3253	$I(1)$
$ln\ S_{it}^{f-CH}$	−2.1669	1.4545	−2.0986	0.8658	−0.3443	$I(1)$
$ln\ S_{it}^{s}$	−1.7103	1.5455	−2.0993	0.8590	1.9686	$I(1)$
$ln\ S_{it}^{f-LP}$	−2.7104	1.2727	−2.1105	0.8541	−3.0447	$I(0)$
$ln\ S_{it}^{f}$	−1.2209	1.0909	−2.1205	0.8427	4.5963	$I(1)$
$(M_{it}/Y_{it})\ ln\ S_{it}^{f}$	−2.4329	1.5455	−2.1040	0.8762	−1.6478	$I(1)$
$(M_{it}/Y_{it})\ ln\ S_{it}^{f-CH}$	−2.4622	1.9545	−2.0929	0.9050	−1.8209	$I(1)$
$(M_{it}/Y_{it})\ ln\ S_{it}^{s}$	−2.3571	1.3182	−2.0941	0.8523	−1.3362	$I(1)$
(M_{it}/Y_{it})	−2.4249	1.3636	−2.0991	0.8575	−1.6501	$I(1)$

(a) cross section average of individual Dickey-Fuller t-statistics
(b) cross section average of individual number of lagged differenced terms in ADF(p_i) regression
(c) cross section average of $E[t_i(p_i, \theta_i)]$
(d) cross section average of $Var[t_i(p_i, \theta_i)]$
(e) the test statistic Ψ_t has a standard normal distribution
(f) test of the null hypothesis of common unit autoregressive root

searcher to worry less about exogeneity considerations.[17] In the absence of a cointegrating relationship, results from a model like (1) in levels would normally be treated with a fair degree of scepticism. The IPS panel unit root tests support Coe and Helpman's original time series analysis, but importantly, they show that simply using LP's alternative foreign R&D stock measure is unlikely to make Coe and Helpman's basic approach more robust.

4. Some Panel Cointegration Models

The next step was to replace Coe and Helpman's original treatment of panel cointegration (i.e., their application of raw panel unit root tests to the estimated residuals from a pooled regression) with the tests recently developed by Pedroni (1997, 1998). Nine model specifications are considered. Models (i) to (iii) are the original Coe-Helpman equations:

$$\ln F_{it} = \alpha_{1i} + \alpha_1^d \ln S_{it}^d + \alpha_1^f \ln S_{it}^{f-CH} ; \quad \text{(i)}$$

$$\ln F_{it} = \alpha_{2i} + \alpha_2^d \ln S_{it}^d + \alpha_2^{G7}\, G7 \ln S_{it}^d + \alpha_2^f \ln S_{it}^{f-CH} ; \quad \text{(ii)}$$

$$\ln F_{it} = \alpha_{3i} + \alpha_3^d \ln S_{it}^d + \alpha_3^{G7}\, G7 \ln S_{it}^d + \alpha_3^f (M_{it}/Y_{it}) \ln S_{it}^{f-CH} ; \quad \text{(iii)}$$

of which (iii) was shown to be mis-specified by LP (the idiosyncratic disturbances have been suppressed since the notational clutter is bad enough already). Model (iv) is LP's preferred version of (iii), using their stationary measure $\ln S_{it}^{f-LP}$ while (v) is another version of (iii) and uses $\ln S_{it}^f$ to correct for *demonstrated* mis-specification, but ignores the *potential* for aggregation bias. As noted in Section 2, however, if we use Coe and Helpman's trade-weighted foreign R&D stock without indexing, then we need to include the aggregate import share as an additional regressor. So, for model (v) we have

$$\ln F_{it} = \alpha_{5i} + \alpha_5^d \ln S_{it}^d + \alpha_5^{G7}\, G7 \ln S_{it}^d + \alpha_5^f (M_{it}/Y_{it}) \ln S_{it}^f + \alpha_5^m (M_{it}/Y_{it}) \ . \quad \text{(v)}$$

Note that while model (v) may still be cointegrated, this is not a possibility for model (iv) because $\ln S_{it}^{f-LP}$ is trend stationary. Models (vi) through (viii) are Keller's simple sum versions of (i) to (iii) where the un-weighted $\ln S_{it}^s$ replaces the weighted sum $\ln S_{it}^f$. As for model (v), model (viii) also includes the aggregate import share:

[17] That is, OLS estimates of a cointegrating vector are super-consistent in the sense of Stock (1987).

$$\ln F_{it} = \alpha_{8i} + \alpha_8^d \ln S_{it}^d + \alpha_8^{G7}\, G7 \ln S_{it}^d + \alpha_8^f (M_{it}/Y_{it}) \ln S_{it}^s + \alpha_8^m (M_{it}/Y_{it}) \,. \quad \text{(viii)}$$

Finally, model (ix) asks if it's just the aggregate import share that is driving all these findings of significant R&D spillovers:

$$\ln F_{it} = \alpha_{9i} + \alpha_9^d \ln S_{it}^d + \alpha_9^{G7}\, G7 \ln S_{it}^d + \alpha_9^m (M_{it}/Y_{it}) \,. \quad \text{(ix)}$$

Model (ix) allows for the possibility that the highly significant coefficients obtained on $(M_{it}/Y_{it}) \ln S_{it}^{f-CH}$ for model (iii) are really driven by (M_{it}/Y_{it}). This possibility is suggested by Keller's finding that it does not seem to matter too much how $\ln S_{it}^f$ is approximated.

Pooled Estimation Results

Table 2 reports the pooled least squares estimates of these nine models with the results from Pedroni's (1997, 1998) panel ADF test at the bottom.[18] Again, these statistics are distributed standard normal with significantly negative statistics indicating rejection of the null hypothesis of no cointegration. Although all but two of these models are capable of supporting more than a single cointegrating vector, a simple test for the existence of cointegration is all that's being performed here. In a panel with a small cross section, Johansen's procedure could conceivably be used to investigate multiple cointegrating relationships more closely, but that option is infeasible with a 22 $\times$ 20 panel.

Five of the nine models are cointegrating, including all of Coe and Helpman's original models and the versions of Keller's simple sum approach that correspond to (i) and (ii). The test statistic for model (iv) is not reported since we know that it cannot be cointegrated. Note that the two specifications derivative of Coe and Helpman's preferred model cannot reject the no cointegration null hypothesis. Excluding (iii) on the grounds of its known misspecification, we are left with two pairs of models that exhibit cointegrating relationships: Coe and Helpman's original models (i) and (ii) and Keller's versions of those same models with $\ln S_{it}^s$ instead of $\ln S_{it}^f$. Before moving on to look at the results from group mean estimation, it's worth briefly comparing these estimates to those obtained by earlier papers. Note that while the point estimates are generally very close to those obtained by Coe and Helpman, Keller and LP (varying occasionally at the second decimal place, more often at the third decimal place), two of the models estimated here

[18]These tests are analogous to the Engle-Granger (1987) approach to testing a no cointegration null in an ordinary time series setting. See the Appendix for some of the mechanics and Pedroni (1997, 1998) for the detail.

TABLE 2. *Pooled TFP Estimation, 21 OECD Countries and Israel over 1971–1990 (Annual Frequency)*

Model Estimation Results (a)	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
$ln\ S^d_{it}$	0.0849	0.0784	0.0744	0.0846	0.0532	0.0263	0.0327	0.0483	0.1079
	(0.0093)	(0.0087)	(0.0076)	(0.0090)	(0.0081)	(0.0109)	(0.0105)	(0.0083)	(0.0068)
$G7\ ln\ S^d_{it}$		0.1289	0.1574	0.1354	0.1550		0.0999	0.1611	0.1527
		(0.0158)	(0.0148)	(0.0160)	(0.0142)		(0.0155)	(0.0141)	(0.0159)
$ln\ S^{f-CH}_{it}$	0.1166	0.0847							
	(0.0162)	(0.0155)							
$(M_{it}/Y_{it})ln\ S^{f-CH}_{it}$			0.3043						
			(0.0361)						
$ln\ S^{f-LP}_{it}$				0.2004					
				(0.0477)					
$(M_{it}/Y_{it})ln\ S^f_{it}$					0.4459				
					(0.0436)				
$ln\ S^s_{it}$						0.2219	0.1770		
						(0.0190)	(0.0194)		
$(M_{it}/Y_{it})ln\ S^s_{it}$								0.4476	
								(0.0427)	
(M_{it}/Y_{it})					−5.3475			−6.3764	0.0855
					(0.5328)			(0.6181)	(0.0491)
$\bar{R}^2$	0.6240	0.6749	0.6981	0.6658	0.7233	0.6814	0.7096	0.7260	0.6541
S.E.E.	0.0484	0.0450	0.0434	0.0457	0.0416	0.0446	0.0426	0.0413	0.0465
Cointegration Results									
Panel ADF statistic (b)	−5.5623	−3.684	−2.0647	—	−0.4272	−4.0641	−2.2203	−0.6846	3.3308
Decision (c)	*CI*	*CI*	*CI*	(d)	retain null	*CI*	*CI*	retain null	retain null

(a) OLS fixed effects estimation (fixed effects dummies unreported) with standard errors in parentheses
(b) panel ADF statistic allows dynamics and cointegrating vector to vary across individuals
(c) test of the null hypothesis of no cointegration (spurious regression)
(d) omitted since equation does not meet the necessary conditions for the existence of a cointegrating relationship.

produce quite different results from those originally reported.[19] The results for model (iv) are qualitatively the same as the estimates obtained by LP, but for each coefficient, slightly higher numbers were obtained. This is probably because, as noted above, static bilateral trade shares were used in this study while LP used time-varying trade shares to compute their S_{it}^{f-LP} measure of the foreign R&D capital stock. Similarly, the results for model (viii) here are qualitatively the same as the estimates obtained by Keller, but for each variable, the estimated coefficient is slightly lower in this paper.

For models (v) and (viii), which use the aggregate import share as an additional regressor, the elasticity of productivity with respect to the import share is calculated as

$$\varepsilon_{F,M}^{j,k} = \alpha_j^f \left(\frac{1}{440} \sum_{i=1}^{22} \sum_{t=1}^{20} \ln S_{it}^k \right) + \alpha_j^m , \qquad j \in \{5, 8\}, k \in \{f, s\} .$$

Since the panel mean of $\ln S_{it}^f$ is about 11.9, we have $\varepsilon_{F,M}^f = 11.9 \times 0.4459 - 5.3475 \approx 0$, which is consistent with the results obtained by Coe and Helpman (1995) and Coe and Hoffmaister (1999). However, the panel mean of $\ln S_{it}^s$ is about 13.3, so we have $\varepsilon_{F,M}^s = 13.3 \times 0.4476 - 6.3764 \approx -0.4$. Neither Coe and Helpman nor Coe and Hoffmaister record such as a calculation for the simple sum measure of foreign R&D, so no direct comparison with their results for this measure can be made.

Group Mean Estimation Results

The pooled results rely on homogeneous panel cointegration theory (they impose common slope coefficients).[20] However, Coe and Helpman originally found that the elasticity of productivity with respect to domestic R&D capital is substantially higher for the G7 countries than for the other industrialized countries in their sample. Also, the elasticity with respect to foreign R&D is typically higher for the smaller countries than it is for the G7. In principle, then, it might be suspected that this slope heterogeneity exists not only between G7 and non-G7 groups but also between individual

[19]Although the standard errors of coefficient estimates from a cointegrating regression are biased (and not even asymptotically normal) they are reported in parentheses for completeness' sake.

[20]In their Appendix B, Coe and Helpman (1995, 884–6) experiment with Equation (1) by including time dummies and by interacting linear trends with each of R&D capital stock measures. In this paper, slope coefficient heterogeneity is introduced across countries, but each country retains a constant slope over time. It is cross section heterogeneity that must be addressed for the purposes of the panel cointegration tests. Time series heterogeneity just adds more "nuisance" parameters.

pairs of countries. If the G7 has a higher domestic R&D elasticity than the non-G7, it might be preferable to let the United States have a different slope coefficient to that of Italy or Canada.[21]

One alternative to pooled estimation is to estimate the model equation separately for each member of the cross section and to analyze the distribution of coefficient estimates: in particular, the mean of the group estimates can be examined. Table 3 presents the results of this exercise for the nine models discussed above. The coefficient reported is the mean of the individual OLS estimates while in parentheses is not the standard error of a point estimate but the standard deviation of the *obtained distribution* of the coefficient estimates. The results of Pedroni's (1997, 1998) group mean panel cointegration test are reported at the bottom of the table. Again, these panel test statistics are distributed (asymptotically) standard normal. It's somewhat reassuring that the decisions from these heterogeneous panel cointegration tests are exactly the same as for the homogeneous tests reported in Table 2.

Two things are striking about these group mean results. First, the mean estimates are generally very different from those obtained from data pooling. For example, in models (i) and (ii) the pooled estimate of the elasticity of TFP with respect to the domestic R&D capital stock is about 0.08 for the non-G7 countries and about 0.20 for the G7 countries. The group mean estimates from models (i) and (ii) are 0.20 to 0.30 for the non-G7 countries and 0.47 for the G7. Even more important is that for models (i) through (iv) the elasticity of TFP with respect to a measure of foreign R&D capital is *negative*. However, keep in mind that the group mean method is consistent (but not usually efficient) for the mean of the slope coefficients *given* that the individual time series are cointegrated. When the group mean cointegration test retains the null—as it does for models (v), (viii) and (ix)—the group mean estimates are generically biased, inconsistent and inefficient, and hence of dubious value. It is only when a panel cointegration relationship can be found that we can put some trust in the obtained group mean estimates.

Second, the distribution of coefficient estimates is generally very broad. In model (ii), a one-standard-deviation band around the 0.47 mean estimate for the G7 domestic R&D elasticity would put that average elasticity somewhere between -0.02 and 0.96. Similarly broad distributions characterize almost all of these mean estimates. These estimation results raise a couple of interesting issues for future research on international R&D spillovers.

[21]The argument in this paragraph draws extensively on comments made by an anonymous referee.

TABLE 3. *Group Mean TFP Estimation, 21 OECD Countries and Israel over 1971–1990 (Annual Frequency)*

Model Estimation Results (a)	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
$ln\ S^d_{it}$	0.2969	0.2161	0.2086	0.2197	−0.0247	0.0994	0.0880	0.0027	0.1598
	(0.3201)	(0.1686)	(0.1683)	(0.2333)	(0.7188)	(0.5519)	(0.6208)	(0.7745)	(0.1229)
$ln\ S^d_{it}$ (*G7 separated*)		0.4701	0.4197	0.4891	0.0228		0.1238	0.0805	0.3634
		(0.4910)	(0.4174)	(0.3950)	(0.4325)		(0.4071)	(0.5716)	(0.3547)
$ln\ S^{f-CH}_{it}$	−0.0640	−0.0640							
	(0.1029)	(0.1029)							
$(M_{it}/Y_{it})\ ln\ S^{f-CH}_{it}$			−0.1528						
			(0.3674)						
$ln\ S^{f-LP}_{it}$				−0.2357					
				(0.3924)					
$(M_{it}/Y_{it})\ ln\ S^f_{it}$					0.3261				
					(0.7949)				
$ln\ S^s_{it}$						0.1490	0.1490		
						(0.7617)	(0.7617)		
$(M_{it}/Y_{it})\ ln\ S^s_{it}$								0.2653	
								(0.8544)	
(M_{it}/Y_{it})					0.1761			0.1987	0.0591
					(0.4146)			(0.4320)	(0.4268)
Cointegration Results									
Group mean statistic (b)	−10.8138	−3.4600	−2.6716	—	−1.1696	−8.2758	−2.1338	−1.2908	2.6211
Decision (c)	*CI*	*CI*	*CI*	(d)	retain null	*CI*	*CI*	retain null	retain null

(a) cross section averages over the individual OLS estimates from time series regression, standard deviation of the individual regression coefficients in parentheses
(b) group mean statistic allows dynamics, cointegrating vectors and autoregressive coefficients to vary across individuals
(c) test of the null hypothesis of no cointegration (spurious regression)
(d) omitted since equation does not meet the necessary conditions for the existence of a cointegrating relationship

Discussion of Estimation Results

Only two specifications come through the panel cointegration tests and alternative estimation approaches looking relatively robust. These are two of the three regressions estimated by Keller (1998) in his demonstration that using a simple-sum definition of a country's foreign R&D capital stock produced higher estimated R&D spillovers than were originally found by Coe and Helpman. Also, the basic model favored by Coe and Helpman, (iii), and its derivatives (iv), (v) and (ix), may be considered the least robust in the sense that the models are either mis-specified (iii), unable to reject the spurious regression null (v, ix), or yield some coefficients that change sign depending on whether pooled or group mean estimation was used (iii, iv). Finally, any model which uses a weighted measure of the foreign R&D capital stock, such as S_{it}^{f-CH}, S_{it}^{f-LP} or S_{it}^{f}, produces a coefficient for the foreign stock that reverses sign when the estimation method is switched. Only models (vi) through (ix) which either use the un-weighted S_{it}^{s} or use no measure of the foreign R&D stock do not exhibit sign reversals. In short, while LP's recommended improvements may be less useful than originally thought, the results from Keller's experiment continue to stand and continue to cast doubt on both the qualitative and quantitative findings reported by Coe and Helpman

In effect, the results in Tables 2 and 3 represent two extreme ways to approach panel estimation with the choice being between (a) complete heterogeneity of slopes and error variances with conditionally consistent but inefficient estimators or (b) complete homogeneity of slopes and error variances but possibly inconsistent estimators.[22] For all the models considered in this paper, the specification is a simple relationship between the levels of variables. Consequently, one possible explanation for the strange group mean coefficient estimates obtained here is the downward finite sample bias that arises from the mis-specified dynamics of a simple levels regression, as documented by Banerjee et al (1986). Since this bias is downward for all individuals, taking cross section averages will not correct for it. A standard way around this finite sample bias is to move to using either single-equation error correction or fully modified OLS approaches.

Although Coe and Helpman originally considered simple error correction models, the ECM approach has not been pursued here since most of the subsequent research has focused on the levels equations. Also, while ECMs would overcome the problem of deriving results purely from static Engle-Granger type regressions, under group mean methods they would

[22]This "choice" is somewhat independent of the treatment of homogeneity for the purposes of integration/cointegration tests since it may be possible to model (and allow for) heterogeneous autocorrelation structures in pre-testing while still using homogeneous estimation methods.

require estimating an infeasibly large number of additional parameters for each cross section member. All is not lost, however, since moving to a conditional ECM (or, equivalently, ARDL) approach would also allow the use of the "intermediate" method of obtaining estimators from dynamic panel data advocated by Pesaran, Shin and Smith (1998). That is, one could use a conditional error correction framework where long-run elasticities are constrained to be the same but where short-run dynamics are allowed to vary over the cross section.

Alternatively, the pooled FM approach proposed by Phillips and Moon (1999a, 1999b) offers a different advantage. As noted above, the group mean approach is valid if and only if the individual time series exhibit cointegration, but Phillips and Moon show that the coefficients obtained from estimating a pooled FM model have an interpretation as a long-run *average* relation irrespective of the existence of individual cointegrating relationships. Consequently, while the estimates presented in Table 3 for the spurious regression models (v, viii and ix) are barely interpretable, the corresponding estimates from a pooled FM model would at least have the advantage of being interpretable as a long-run average statistical relation *even in the absence of individual cointegrating relationships.*[23] Future research on international R&D spillovers may benefit from exploiting the FM approach.[24]

5. Conclusion

Recently, Coe and Helpman (1995) estimated the relationship between levels of total factor productivity and levels of domestic and foreign R&D capital stocks. Two issues arising from Coe and Helpman's use of panel data are addressed in this paper: the application of raw unit root tests to obtained residuals in a panel setting, and the use of pooled estimation techniques.

Pedroni (1997) has shown that when applying raw panel unit root tests to obtained residuals, care has to be taken with specifying the homogeneity properties of the panels. The first result of this paper is to show that Coe and Helpman's conclusions about the integration/cointegration structure of their models are unchanged. The same conclusions in favor of cointegrating equations are reached irrespective of whether panel (Levin-Lin) or group

[23]This phenomenon arises because of the nature of the long-run variance-covariance matrix of the panel. See Phillips and Moon (1999b).

[24]One objection to this recommendation might be that this strategy would take the international R&D spillovers literature (even) further from its foundations in Grossman and Helpman (1991) and push the literature towards answering questions posed by econometric methodology rather than economic intuition and theory. Of course, this paper is already largely guilty of that charge.

(IPS) methods are used. No longer do researchers following Coe and Helpman have to defend the use of a simple levels specification like Equation (1) by *purely* appealing to a priori intuition about the plausibility of coefficient estimates.

However, Coe and Helpman's results are sensitive to the estimation method used. Given the likely economic relevance of a heterogenous slope coefficient specification, at least some attention should also be paid to alternative group mean estimation results. The second result of this paper is to show that only Keller's models come through panel cointegration tests and group mean estimation approaches looking relatively robust. Models derivative of Coe and Helpman's preferred specification, Equation (1), including the alternative suggested by LP, come out either mis-specified, unable to reject a spurious regression null, or yield coefficients that change sign depending on whether pooled or group mean estimation was used.

The problems identified in the second half of this paper suggest that despite the findings of panel cointegration, a more robust approach to estimating international R&D spillovers might better be developed from either panel ARDL or pooled FM methods. Although the ARDL methods advocated by Pesaran, Shin and Smith (1998) do not allow for the same degree of cross section heterogeneity allowed for in this paper, these methods assist in overcoming bias in the estimates of long-run parameters (due to the mis-specified dynamics of a static levels regression) while still permitting a greater degree of cross section heterogeneity than allowed for in Coe and Helpman's brief use of an error correction framework. Alternatively, the pooled FM methods proposed by Phillips and Moon (1999a, 1999b) would allow pooled estimates to be interpreted as a measure of a long-run average statistical relationship even in the absence of formal panel cointegration.

Received: April 1999
Final version: April 2000

References

Banerjee, Anindya, Juan J. Dolado, David F. Hendry, and Gregor W. Smith. "Exploring Equilibrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence." *Oxford Bulletin of Economics and Statistics* 48 (1986): 253–77.

Coe, David T., and Alexander W. Hoffmaister. "Are There International R&D Spillovers among Randomly Matched Trade Partners? A Response to Keller." International Monetary Fund, Working Paper, 99/18, 1999.

Coe, David T., and Elhanan Helpman. "International R&D Spillovers." *European Economic Review* 39 (1995): 859–87.

Coe, David T., Elhanan Helpman and Alexander W. Hoffmaister. "North-South R&D Spillovers." *Economic Journal* 107 (1997): 134–49.

Engle, Robert F., and Clive W. J. Granger. "Co-integration and Error Correction: Representation, Estimation and Testing." *Econometrica* 55 (1987): 251–76.

Grossman, Gene, and Elhanan Helpman. *Innovation and Growth in the Global Economy.* Cambridge, MA: MIT Press, 1991.

Im, Kyung So, M. Hashem Pesaran and Yongcheol Shin. "Testing for Unit Roots in Heterogeneous Panels." Manuscript, University of Cambridge, 1997

Kao, Chihwa. "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data." *Journal of Econometrics* 90 (1999): 1–44.

Keller, Wolfgang. "Are International R&D Spillovers Trade Related? Analyzing Spillovers among Randomly-Matched Trade Partners." *European Economic Review* 42 (1998): 1469–81.

Keller, Wolfgang. "How Trade Patterns and Technology Flows Affect Productivity Growth." *National Bureau of Economic Research,* Working Paper, 6990, 1999.

Kremers, Jereon J. M., Neil R. Ericsson and Juan J. Dolado. "The Power of Cointegration Tests." *Oxford Bulletin of Economics and Statistics* 54 (1992): 325–48.

Levin, Andrew, and Chien-Fu Lin. "Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties." University of California, San Diego, Discussion Paper 92-23, 1992.

Levin, Andrew, and Chien-Fu Lin. "Unit Root Tests in Panel Data: New Results." University of California, San Diego, Discussion Paper 93-56, 1993.

Lichtenberg, Frank, and Bruno van Pottelsberghe de la Potterie. "International R&D Spillovers: A Comment." *European Economic Review* 42 (1998): 1483–91.

Pedroni, Peter. "Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis: New Results." Manuscript, Indiana University, 1997.

———. "Approximate Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors." Manuscript, Indiana University, 1998.

Pesaran, M. Hashem, and Ron Smith. "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels." *Journal of Econometrics* 68 (1995): 79–113.

Pesaran, M. Hashem, Yongcheol Shin and Ron Smith. "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels." Manuscript, 1999. *Journal of the American Statistical Association,* forthcoming.

Phillips, Peter C. B., and Bruce E. Hansen. "Statistical Inference in Instrumental Variables Regression with $I(1)$ Processes." *Review of Economic Studies* 57 (1990): 99–125.
Phillips, Peter C. B., and Hyungsik R. Moon. "Nonstationary Panel Data Analysis: An Overview of Some Recent Developments." Cowles Discussion Paper, 1999a.
———. "Linear Regression Limit Theory for Nonstationary Panel Data." *Econometrica* 67 (1999b): 1057–1112.
Quah, Danny T. "Exploiting Cross-Section Variation for Unit Root Inference in Dynamic Data." *Economics Letters* 44 (1994): 9–19.
Stock, James H. "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors." *Econometrica* 55 (1987): 1035–56.

Appendix

The discussion in this appendix draws heavily on the material presented in Im, Pesaran and Shin (1997, 2–10) and Pedroni (1998, 8–9), which should be consulted for further details.

1. Im, Pesaran and Shin's (1997) t-bar panel unit root test for autocorrelated series. Consider a panel of data on the variable y_{it}. Run augmented Dickey-Fuller regressions of the form

$$\Delta y_{it} = \gamma_i + \beta_i y_{i,t-1} + \sum_{L=1}^{p_i} \theta_{iL} \Delta y_{i,t-L} + \epsilon_{it}, \qquad \text{for each } i \in N,$$

where the lag truncation order for each individual, p_i, is the minimum necessary to purge autocorrelation from ϵ_{it}. We need $\epsilon_{it} \sim iidN(0, \sigma_i^2)$ and N, $T \to \infty$ at a rate such that $N/T \to k$, a finite positive constant. Then obtain the t_i-statistic for testing the null hypothesis $\beta_i = 0$. The average over N of these $t_i(p_i, \theta_i)$ is the t-bar reported in Table 1. IPS show that, if the t_i have finite second moments, the following modified test statistic

$$\Psi_{\bar{t}} = \frac{\sqrt{N}\,\{\bar{t}(p, \theta) - N^{-1} \sum_i E[t_i(p_i, \theta_i)]\}}{\sqrt{N^{-1} \sum_i Var\,[t_i(p_i, \theta_i)]}},$$

has a (weakly) standard normal distribution. In Table 1, the column "mean adjustment" is the value for $N^{-1} \sum_i E[t_i(p_i, \theta_i)]$ using the individual $E[t_i(p_i, \theta_i)]$ reported by IPS. Similarly, the variance adjustment column reports the $N^{-1} \sum_i Var[t_i(p_i, \theta_i)]$. The group mean statistic is $\Psi_{\bar{t}}$ itself.

2. Pedroni's (1997, 1998) panel ADF and group mean cointegration tests. First estimate the appropriate levels regression, say (1), to obtain an

estimated residual series ϵ_{it}. Second, difference the levels equation once and obtain the residuals, v_{it}. Obtain an estimate of the long-run variance of v_{it}; call it $\sigma^2_{v(i)}$, using a kernel estimator (choosing sample covariance weights appropriately). Third, run ADF like regressions

$$\Delta\epsilon_{it} = \rho_i\epsilon_{i,t-1} + \sum_{L=1}^{p_i} \theta_{iL}\Delta\epsilon_{i,t-L} + u_{it} ,$$

and compute the variance of the residual series; call it $\sigma^2_{u(i)}$. Compute the panel ADF test statistic as

$$Z^* = \left(\bar{\sigma}^2 \sum_{i=1}^{N}\sum_{t=1}^{T} \sigma^{-2}_{v(i)}\epsilon^2_{i,t-1}\right)^{-1/2} \sum_{i=1}^{N}\sum_{t=1}^{T} \sigma^{-2}_{v(i)}\epsilon_{i,t-1}\Delta\epsilon_{it} , \qquad \text{(A1)}$$

where $\bar{\sigma}^2 = N^{-1}\sum_i \sigma^2_{u(i)}$. Compute the group mean test statistic as

$$N^{-1/2}\tilde{Z}^* = N^{-1/2}\left(\sum_{i=1}^{N}\left(\sum_{t=1}^{T} \sigma^2_{u(i)}\epsilon^2_{i,t-1}\right)^{-1/2}\right)\sum_{t=1}^{T} \epsilon_{i,t-1}\Delta\epsilon_{it} . \qquad \text{(A2)}$$

Then one can obtain the panel cointegration test statistics reported in Tables 2 and 3 by applying the mean adjustment, μ, and variance adjustment, v, reported by Pedroni (1998, 13) to the test statistics from (A1) or (A2). If Ψ_{NT} is the standardized statistic from (A1) or (A2), then the panel cointegration test statistic

$$\frac{\Psi_{NT} - \mu\sqrt{N}}{\sqrt{v}} ,$$

has a standard normal distribution.