

NOTES

DETERMINANTS OF INDIA'S SOFTWARE EXPORTS AND GOODS EXPORTS

P. K. M. Tharakan, Ilke Van Beveren, and Tom Van Ourti*

Abstract—Recent export experience of some large, emerging economies has raised important questions about the trade determinants of the modern-services-driven sectors and the goods-production-driven sectors. In our empirical analysis of the determinants of Indian exports of software services and of the total Indian goods exports, we raise the following questions: How (dis)similar is the performance of the Indian exports of software from the determinants of India's total exports of goods? Are such differences significant? Is the pattern of the performance of the determinants stable over time? Our findings concerning the effects of size, distance, linguistic connections, and trade-facilitating networks enable us to make some important inferences of policy relevance.

I. Introduction

In this study we carry out an empirical analysis of the determinants of Indian exports of software and of India's total goods exports. Our paper is organized as follows: Section II selectively reviews some of the recent gravity model applications to nongoods transactions. The empirical approach and results obtained are discussed in section III. Section IV concludes.

The impressive performance of the Indian software industry is well documented and analyzed.¹ The sector is outward oriented, with exports amounting to US\$9.55 billion in financial year 2002–2003, equaling 18% of India's total goods exports (Nasscom, 2004). Export growth rates averaged above 50% per year in the 1990s. Software exports are service- rather than product-oriented.²

The reasons cited by researchers and software business sources³ for the remarkable success of Indian software exports include: the rapidly growing world demand for software, the cost difference between India and the outsourcing countries in employing software professionals, knowledge of the English language by Indian software professionals, and the possibility of interacting with clients at low cost either through professionals' mobility (on-site services) or through high-speed data-com links (offshore services). The positive role played by people of Indian origin in various countries in facilitating software trade as well as trade in general with India is also sometimes mentioned.

Received for publication January 2, 2003. Revision accepted for publication December 9, 2004.

* University of Antwerp, University of Antwerp and Lessius Hogeschool, and University of Antwerp and Erasmus University Rotterdam, respectively.

We are indebted to a large number of people for their help and encouragement in carrying out this study. In India, IT professionals from Infosys Ltd., N.I.I.T., and Tata Consultancy Services provided us valuable information. K. J. Joseph, Thomas Abraham, and C. Veeramani were of help in assisting to obtain the necessary data. V. Ginsburg provided valuable advice on the econometric part of the study. We also thank an anonymous referee for providing useful comments, and an editor of this review for some specific suggestions.

¹ Arora, et al. (2001), Banerjee and Duflo (2000), and Joseph and Harilal (2001) are among the useful sources.

² Of the total software exports of India in FY 2001–2002, 95.89% consisted of on-site services and offshore services. Only 4.11% consisted of products and packages.

³ Based on discussions with I.T. professionals from Infosys Ltd., N.I.I.T., and Tata Consultancy Services as well as on the studies on the Indian software industry cited earlier.

II. Brief Overview of Literature

The capacity of the gravity model to empirically explain bilateral goods trade has been widely recognized since the early publications in this field by Tinbergen (1962), Pöyhönen (1963), and others. Yet it lacked a sound theoretical foundation. Anderson (1979) was the first to furnish a solid theoretical foundation for gravity models. More recently, Anderson and van Wincoop (2003) have developed a method that consistently and efficiently estimates a theoretical gravity equation. Their basic insight is that multilateral resistance should be allowed for appropriately while estimating gravity models.

Other important recent results include those yielded by the application of gravity models to international transactions other than those in goods. Examples include: foreign direct investment flows and stocks (Brenton, Di Mauro, & Lücke, 1999); stock market correlations (Flavin, Hurley, & Rousseau, 2002); telephone calls and communication barriers (Matthes, 1994); and financial assets (Portes, Rey, & Oh, 2001). Somewhat surprisingly, in all of the studies cited above, a significantly negative coefficient was obtained for the distance variable. Portes and Rey (2002) suggest that distance is a proxy not only for transportation costs, but also for other transaction and information costs (*informational asymmetries*) associated with long-distance exchange.

III. Empirical Model and Results

It is important to keep in mind that the gravity model which we implement is of a partial equilibrium nature, dealing with the exports of software of one country (India) to all the other countries, and not the bilateral trade between all countries of the world. Hence by its very nature, our exercise will be able to implement only part of the innovations made by Anderson and van Wincoop (2003). We come back to this issue in section III D on sensitivity analyses.

A. Gravity Model Specification

The gravity model, in its basic form, relates bilateral trade directly to the economic masses (GDP) of the trading partners and to the distance between them. Examples of other variables included in the model are population, adjacency, membership in preferential trading arrangements, and prevalence of use of the same language. More recently the role played by trade-facilitating networks has been stressed (Gould, 1994; Rauch & Trindade, 2002).

The following gravity model specifications are used to evaluate the determinants of India's bilateral software exports⁴ (SE_{ji}) and total goods exports (TE_{ji}):

⁴ We prefer Nasscom data (to data from the Department of Electronics, the Reserve Bank of India, and the Centre for Monitoring the Indian Economy) because they provide a countrywise breakdown of exports for each year. For a critical appraisal of the Indian software trade data, see Joseph and Harilal (2001). In order to cross-check the accuracy of the Nasscom data, we contacted 51 national IT associations, 45 national statistical offices, and 9 international data sources: none of them were able to provide import figures for software from India. However, aggregate exports of software available from the International Data Corporation in

$$\begin{aligned} \ln(SE_{jt}) = & \beta_0 + \beta_1 \ln(GDP_{jt}) + \beta_2 \ln(POP_{jt}) \\ & + \beta_3 \ln(D_{\text{India},j}) + \beta_4 \ln(REM_{jt}) + \beta_5 LANG_j \\ & + \beta_6 \ln(PIO_{jt}) + \sum_{t=1997}^{2001} \beta_7 TIME_t + u_{jt}, \end{aligned} \quad (1a)$$

$$\begin{aligned} \ln(TE_{jt}) = & \gamma_0 + \gamma_1 \ln(GDP_{jt}) + \gamma_2 \ln(POP_{jt}) \\ & + \gamma_3 \ln(D_{\text{India},j}) + \gamma_4 \ln(REM_{jt}) + \gamma_5 LANG_j \\ & + \gamma_6 \ln(PIO_{jt}) + \gamma_7 \ln(DUTIES_j) \\ & + \gamma_8 \ln(GSP_j) + \sum_{t=1997}^{2001} \gamma_9 TIME_t + v_{jt}, \end{aligned} \quad (1b)$$

where j refers to country, t to time, $TIME_t$ stands for a time dummy, and u_{jt} and v_{jt} are stochastic random components that are normally distributed with mean 0 and variances σ_u^2 and σ_v^2 .⁵ Further, GDP_{jt} represents the economic size of the importing country (or its relative difference in size from India) and is expected to have a positive coefficient in both equations. The sign of population size (POP_{jt}) cannot be predicted a priori since its coefficient can reflect economies of scale in production or specialization through a greater division of labor (Gould, 1994).⁶

The distance variable ($D_{\text{India},j}$) consists of the great circle distance (GCD) between the capitals of India and country j .⁷ In addition to bilateral distance, a remoteness variable⁸ (REM_{jt}) enters the equation. Following Helliwell (1997) and Rauch and Trindade (2002), REM_{jt} is defined as the average weighted distance of j from all its other potential trading partners, where GDP_{jt}^{-1} is used as weight, that is, $REM_{jt} = \sum_{m \neq \text{India}} D_{j,m} / GDP_{m,t}$. For software services, a significant coefficient is expected for both variables if information asymmetries, as proxied by distance (see section II), matter. For total goods exports the distance variable should yield the traditional negative sign, whereas the remoteness term is expected to yield a positive and significant result.

Because of the dominance of the English language in software services markets and because India has the second largest group of English-speaking scientific professionals in the world (Nasscom, 2001, p. 43), software export transactions with countries where English is a national language are likely to be facilitated. Hence for software exports a positive sign for the English language dummy ($LANG_j$) is expected, but we are less sure about such an outcome for India's geographically more diverse goods exports.

Gould (1994) and Rauch and Trindade (2002) have emphasized the trade-facilitating role played by networks through the provision of

market information and matching services. Anecdotal evidence suggests that people of Indian origin might have been active in facilitating bilateral software transactions, particularly those related to the United States (see Arora et al., 2001; Saxenian, 1999). Following Rauch and Trindade (2002), we define our variable PIO_{jt} as the share of people of Indian origin in country j 's total population. We expect a positive effect in both equations.

We now turn to the two variables specific to India's total goods exports. Because most favored nation (MFN) duties⁹ ($DUTIES_j$) are low after successive rounds of tariff reductions, the duties need not be an important obstacle for India's total goods exports. We expect a positive coefficient for the dummy that indicates whether India's total goods exports are eligible for Generalized System of Preference tariff reductions (GSP_j).

Although preferential trading arrangements and adjacency have often figured as explanatory variables in gravity models, neither is relevant here. The South Asian Association of Regional Cooperation (SAARC), of which India is a member, has been particularly unsuccessful in becoming operational. India's trade with some of the adjacent countries is troubled by political tensions. The value of trade with them is also small.¹⁰

B. Econometric Issues

We estimate a modified gravity model in which a *minimum threshold value* α has to be reached before exports actually occur (see Eaton & Tamura, 1994) and Rauch & Trindade, 2002). If $\alpha = 0$, the traditional tobit model is nested within this threshold tobit model.

Because we are primarily concerned with the question of *differing* performance of the regressors in both specifications, we assess the statistical significance and the magnitude of the differences. We calculate t-tests by pooling the data of total and software exports and by using dummies for software. Due to the large number of zero observations for software exports (as compared to total exports) and due to the nonlinearity of the threshold tobit model, we present elasticities to make inferences on the magnitude of the effects. For the continuous regressors (x_{jt}), we present two elasticities of exports (E_{jt}) (evaluated at the mean of the regressors): (i) an elasticity restricted to the sample of positive observations ($\epsilon_x^E|_{E>0} = [\partial E(E_{jt}|E_{jt} > 0; x_{jt}) / \partial x_{jt}] [x_{jt} / E(E_{jt}|E_{jt} > 0; x_{jt})]$) and (ii) an elasticity for the entire sample ($\epsilon_x^E = [\partial E(E_{jt}|x_{jt}) / \partial x_{jt}] [x_{jt} / E(E_{jt}|x_{jt})]$). For the dummy regressors ($LANG_j$ and GSP_j), we calculate the percentage increase (evaluated at the mean of the regressors) in the conditional mean due to a change in the dummy from 0 to 1.

We have tested for multiplicative heteroskedasticity: the LR statistics did not reject the null hypothesis of homoskedasticity. Nevertheless, in order to assure conservative statistical inference, we use robust standard errors and correct the standard errors for clustering at the country level. Hence the statistical significance of our results should be interpreted as the outcome of a very stringent test.¹¹

C. Empirical Results

Table 1 shows the coefficient estimates for equations (1a) and (1b). Table 2 displays the t -tests, and table 3 presents the elasticities. The economic size of the importing countries (GDP_{jt}) has a highly

India for three years in our sample are of comparable magnitude to the Nasscom data.

⁵ Because India appears in each pair of countries, it makes no difference whether we multiply GDP_{jt} and POP_{jt} by their respective value for India, or whether we simply use the value of those variables pertaining to the j 's only. We refer to the appendix for a discussion of the data sources. We exclude countries for which data on GDP, PIO or duties (for total goods exports) were not available. The software (total goods) export values in the sample account for 98% (88%) of exports.

⁶ Although including GDP per capita is mathematically equivalent to including population (see Frankel, 1997, p. 57), population is preferred due to its lower correlation with GDP, duties, and GSP.

⁷ For the GCD calculation, we refer to Schumacher (2001, p. 9).

⁸ We are thankful to an anonymous referee for pointing out the importance of correcting for remoteness even when there is only one exporter (India) and many j 's.

⁹ Average bound duties for all products, defined as $(1 + \text{duty})/100$.

¹⁰ Only two of the countries from SAARC and three from the adjacent category even figure in either sample.

¹¹ To illustrate this, we point out that regressing without clustering halves the standard errors of the majority of the coefficients.

TABLE 1.—THRESHOLD TOBIT ESTIMATION OF POOLED EXPORTS (1997–2001)

Dependent Variable	$\ln(SE_{jt})$	$\ln(TE_{jt})$
α	3.28* (0.71)	1.66† (0.73)
σ	0.97* (0.11)	0.92* (0.09)
Constant	37.24 (28.74)	-7.27 (39.00)
$TIME_{1998}$	0.17 (0.16)	0.12 (0.17)
$TIME_{1999}$	0.08 (0.12)	0.13 (0.11)
$TIME_{2000}$	0.31† (0.14)	0.15 (0.09)
$TIME_{2001}$	0.09 (0.15)	0.26* (0.09)
$\ln(GDP_{jt})$	0.83* (0.09)	0.54* (0.10)
$\ln(POP_{jt})$	-0.19† (0.10)	0.33* (0.08)
$\ln(D_{India,j})$	0.20 (0.47)	-1.02‡ (0.62)
$\ln(REM_{jt})$	-6.62 (4.80)	1.18 (6.53)
$LANG_j$	1.13* (0.29)	0.15 (0.22)
$\ln(PIO_{jt})$	0.13* (0.04)	0.13* (0.03)
$\ln(DUTIES_j)$	—	-0.06 (0.95)
GSP_j	—	0.06 (0.30)
N	640	510
$N _{Exports>0}$	266	485
Log likelihood	-1,212.70	-2,663.37

Maximum likelihood estimation of threshold tobit model.

Robust standard errors, corrected for clustering at the country level, in parentheses.

* Significant at 1% level; † significant at 5% level; ‡ significant at 10% level.

significant and positive coefficient in both regressions, but the null hypothesis of equality of the coefficients is rejected. As the elasticities show, the import demand generated from the economically large countries exerts a greater effect on India's software exports sector than on its total goods exports. POP_{jt} yields contrasting results for the two regressions. Its coefficient is significantly negative for software exports, but it is positive and significant for total exports. The equality test clearly rejects the null hypothesis.

The distance variable does not yield a significant coefficient for India's software exports. This result suggests that distance is not an important trade resistance factor for software exports, in sharp contrast with earlier results obtained for other international transactions that are in essence "weightless" (see section II). The null hypothesis of equal performance across equations is clearly rejected. The remoteness term does not yield any significant result for either type of exports.

The English language dummy yields a positive and highly significant coefficient for software exports, but for total goods exports the

TABLE 3.—MAGNITUDES OF EFFECTS OF REGRESSORS: ELASTICITIES

Variable	$\epsilon_x^E _{E>0}$ (Positive Observations)		ϵ_x^E (All Observations)	
	$\ln(SE_{jt})$	$\ln(TE_{jt})$	$\ln(SE_{jt})$	$\ln(TE_{jt})$
$\ln(GDP_{jt})$	0.86	0.54	0.90	0.54
$\ln(POP_{jt})$	-0.20	0.33	-0.21	0.33
$\ln(D_{India,j})$	0.21	-1.03	0.22	-1.03
$\ln(REM_{jt})$	-6.89	1.19	-7.22	1.19
$LANG_j$	1.58*	0.16*	1.58*	0.16*
$\ln(PIO_{jt})$	0.14	0.13	0.15	0.13
$\ln(DUTIES_j)$		-0.06		-0.06
GSP_j		0.06*		0.06*

* Percentage increase.

coefficient is insignificant. The equality test in table 2 confirms the different performance of the coefficients in the two regressions. Given the importance of English in software development compared to the heterogeneous nature of India's total goods exports, this is a reasonable outcome. The share of the people of Indian origin (PIO_{jt}) in the importing country's population yields a positive and highly significant result for both specifications, but as table 2 shows, in this case we cannot reject the null hypothesis of equality of the coefficients. This outcome suggests that though network connections positively influence both software exports and total goods exports, they are not the likely channel through which Indian exporters of software overcome the informational asymmetries caused by distance. Average MFN duties and the GSP dummy, which are relevant only for India's total goods exports, do not yield significant results.

D. Sensitivity Analysis¹²

First, we checked for outliers. We estimated equations (1a) and (1b) excluding all observations on Indian exports to the United Kingdom and the United States, the two largest importers of Indian software. The resulting coefficients and equality tests are qualitatively similar to the ones in tables 1 and 2. The largest differences were found for the coefficients of $\ln(GDP_{jt})$ and $LANG_j$ in the software equation; i.e. they change from 0.83 to 0.67 and from 1.13 to 0.84, respectively. Second, we used OLS to estimate equations (1a) and (1b) without the zero-valued observations. The pattern of the coefficients is again very similar to the pattern resulting from the threshold tobit estimates. No coefficient that was previously significant becomes insignificant or vice versa. Third, we estimated equations (1a) and (1b) year by year. The results of the equality tests, reported in table 4, are in line with the t -tests in table 2, and seem to indicate that the difference in performance does not decrease over time (except for the language variable).

As we have seen before, Anderson and Van Wincoop (2003) have developed a general equilibrium theory for the gravity model (see section III). Although adding country fixed effects will yield consis-

¹² Further details on all econometric analyses can be obtained from the authors.

TABLE 2.— t -TESTS FOR THE NULL HYPOTHESIS THAT EFFECTS OF REGRESSORS ARE EQUAL FOR SOFTWARE AND TOTAL EXPORTS

Variable:	$\ln(GDP_{jt})$	$\ln(POP_{jt})$	$\ln(D_{India,j})$	$\ln(REM_{jt})$	$LANG_j$	$\ln(PIO_{jt})$
	2.56†	-4.27*	2.04†	-1.35	3.24*	0.24

Equality test (robust standard errors corrected for clustering at the country level) performed by pooling all observations and using dummies for software.

$N = 1,150$ (640 observations for SE_{jt} and 510 observations for TE_{jt}).

* Significant at 1% level; † significant at 5% level.

TABLE 4.—*t*-TEST FOR THE NULL HYPOTHESIS THAT EFFECTS OF REGRESSORS ARE EQUAL FOR SOFTWARE AND TOTAL EXPORTS: YEAR BY YEAR

Variable	$\ln(GDP_{jt})$	$\ln(POP_{jt})$	$\ln(D_{India,j})$	$\ln(REM_{jt})$	$LANG_j$	$\ln(PIO_{jt})$
1997	0.33	-2.41†	0.53	0.62	3.28*	-0.20
1998	1.64‡	-2.76*	2.47†	-0.79	3.87*	0.82
1999	1.82‡	-3.12*	2.48†	-1.51	2.77*	0.40
2000	2.85*	-4.03*	2.17†	-2.77*	2.11†	0.08
2001	4.19*	-4.85*	2.81*	-2.90*	1.38	0.90

Equality test (robust standard errors corrected for clustering at the country level) performed by pooling all observations and using dummies for software

$N = 230$ per year (128 observations for SE_{jt} and 102 observations for TE_{jt}).

* Significant at 1% level; † significant at 5% level; ‡ significant at 10% level.

tent estimates, it is rather uninformative in our case, for all time-invariant regressors ($D_{India,j}$, $LANG_j$, $DUTIES_j$, and GSP_j) are absorbed in the fixed effect (see Baltagi, Egger, & Pfaffermayr, 2003; Girma & Yu, 2002). Note also that our study deals only with exports from India to other countries. Moreover, of all time-varying regressors, only GDP_{jt} is likely to have sufficient time variation. In spite of these limitations, we carried out a fixed effects regression on the positive observations, with and without the remoteness term.¹³ The resulting coefficients for $\ln(GDP_{jt})$ are very similar to those resulting from the OLS regressions on the positive observations, with only the traditionally unpredictable population and the population-influenced PIO-variable showing changes.¹⁴

Given the overall stability which the results have shown, we conclude that the estimates in table 1 are robust to the sensitivity analyses.

IV. Conclusions

The sharp difference in the effects of the determinants on the two flows of trade is clearly confirmed by the results of the equality tests reported in table 2. In fact among the significant result-yielding explanatory variables (table 1) common to the two equations (1a) and (1b), it is only in the case of people of Indian origin that the null hypothesis of equality of the coefficients cannot be rejected. The interesting finding that distance does not have a significant effect on Indian software exports is in contrast with the results obtained for other “weightless” transactions (see section II) and suggests that the Indian software exporters might have overcome the informational asymmetry problem.

Both knowledge of the English language and the ability to tap into network connections are important determinants of the Indian exports of software. This finding too has important implications for the new wave of software exporters, in that some of these countries can also benefit from a working knowledge of English and have the potential for developing trade-facilitating contacts through the presence of earlier waves of emigrants to other countries. It will be those countries that are able to make effective use of a combination of these advantages that will emerge as the most successful suppliers of software services to high-income markets.

Export-oriented service sectors have become highly important in some emerging economies like India. Our results suggest that from a gravitational point of view, such sectors, especially when they have relatively high technology content as in the case of software, might have the possibility to power export growth more strongly than export-oriented sectors product. This raises some intriguing possibilities. For example, although at present China’s goods-production

export oriented economy is performing far more powerfully than that of India, the latter’s modern-services export-oriented economy might help narrow the gap between the two more swiftly than expected.

REFERENCES

- Anderson, James E., “A Theoretical Foundation for the Gravity Equation,” *The American Economic Review* 69:1 (1979), 106–116.
- Anderson, James E., and Eric Van Wincoop, “Gravity with Gravitas: A Solution to the Border Puzzle,” *The American Economic Review* 93:1 (2003), 170–192.
- Arora, Ashish, V. S. Arunachalam, Jai Asundi, and Ronald Fernandes, “The Indian Software Services Industry,” *Research Policy* 30:8 (2001), 1267–1287.
- Baltagi, Badi H., Peter Egger, and Michael Pfaffermayr, “A Generalized Design for Bilateral Trade Flow Models,” *Economics Letters* 80 (2003), 391–397.
- Banerjee, Abhijit V., and Esther Duflo, “Reputation Effects and the Limits of Contracting: a Study of the Indian Software Industry,” *Quarterly Journal of Economics* 115:3 (2000), 989–1017.
- Brenton, P., F. DiMauro, and M. Lücke, “Economic Integration and Foreign Direct Investment,” *Empirica* 26 (1999), 95–121.
- Eaton, J., and A. Tamura, “Bilateralism and Regionalism in Japanese and US Trade and Foreign Direct Investment Patterns,” *Journal of the Japanese and International Economies* 8 (1994), 478–510.
- Flavin, Thomas J., Margaret J. Hurley, and Fabrice Rousseau, “Explaining Stock Market Correlation: A Gravity Model Approach,” *The Manchester School Supplement* 1463:6786 (2002), 87–106.
- Frankel, Jeffrey A., *Regional Trading Blocs in the World Economic System* (Washington DC: Institute for International Economics, 1997).
- Girma, Sourafel and Zhihao Yu, “The Link between Immigration and Trade: Evidence from the United Kingdom,” *Weltwirtschaftliches Archiv* 138:1 (2002), 115–130.
- Gould, David M., “Immigrant Links to the Home Country: Empirical Implications for U.S. Bilateral Trade Flows,” this REVIEW, 76:2 (1994), 302–316.
- Helliwell, John F., “National Borders, Trade and Migration,” *Pacific Economic Review* 2:3 (1997), 165–185.
- Joseph, K.J. and K.N. Harilal, “Structure and Growth of India’s IT Exports: Implications of an Export-Oriented Growth Strategy,” *Economic and Political Weekly* 36:34 (2001), 3263–3270.
- Matthes, Norbert, “Allocation of Mobile Communication Flows: From Micro-Economic Demand Theory to a Gravity Model,” *The Annals of Regional Science* 28 (1994), 395–409.
- Nasscom, *The IT Software and Services Industry in India: Strategic Review 2000* (New Delhi: National Association of Software and Service Companies, 2001)
- , *Nasscom FY 2003 Software and Service Industry Performance*, (National Association of Software and Service Companies, 2004), www.nasscom.org.
- Portes, Richard, and Hélène Rey, “The Determinants of Cross-Border Equity Transaction Flows,” National Bureau of Economic Research working paper no. 7336 (1999). An updated version (2002) is downloadable at <http://www.princeton.edu/~hrey/research.htm> under the heading “Working papers.”
- Portes, Richard, Hélène Rey, and Yonghyup Oh, “Information and Capital Flows: The determinants of Transactions in Financial Assets,” *European Economic Review* 45:4–6 (2001), 783–796.

¹³ As suggested by a referee.

¹⁴ Excluding remoteness from the fixed effects regression does not affect the results.

- Pöyhönen, Pentti, "A Tentative Model for the Volume of Trade between Countries," *Weltwirtschaftliches Archiv* 90:1 (1963), 93–100.
- Rauch, James E. and Vitor Trindade, "Ethnic Chinese Networks in International Trade," in *REVIEW*, 84:1 (2002), 116–130.
- Saxenian, A., *Silicon Valley's New Immigrant Entrepreneurs* (San Francisco: Public Policy Institute of California, 1999).

- Schumacher, Dieter, "Market Size and Factor Endowment: Explaining Comparative Advantage in Bilateral Trade by Differences in Income and Per Capita Income," Deutsches Institut für Wirtschaftsforschung discussion paper no. 259 (2001).
- Tinbergen, Jan, *Shaping the World Economy* (New York: The Twentieth Century Fund, 1962).

APPENDIX

TABLE A1.—DATA SOURCES AND DESCRIPTION

Description	Data Source
Bilateral Indian software exports (U.S.\$million)	National Association of Software and Service Companies (Nasscom): <i>The I.T. Software and Services Industry in India: Strategic Review</i> , various editions.
Bilateral Indian goods exports (U.S.\$million)	International Monetary Fund: <i>Direction of Trade Statistics Yearbook</i> (2002).
GDP* (U.S.\$million) and population	World Bank: <i>World Development Indicators Database</i> , CD-ROM.
Distance (nautical miles)	Latitudes and longitudes of different cities: <i>The Getty Thesaurus of Geographic Names</i> (www.getty.edu). Great-circle distance calculation: see Schumacher (2001).
Remoteness	See the sources for the distance and GDP variables indicated above.
Language	<i>Ethnologue: Languages of the World</i> , 14th ed. (SIL International, www.ethnologue.com).
People of Indian origin	<i>Report of the High Level Committee on the Indian Diaspora</i> , Ministry of External Affairs, Government of India (www.indiandiaspora.nic.in).
GSP	UNCTAD, GSP Web site (www.unctad.org/gsp)
Duties	WTO: Negotiating Group on Market Access (Doc. no. TN/MA/S/4/Rev.1, 2002).

* For some countries alternative sources were consulted to obtain data on GDP.

R&D AND TECHNOLOGY TRANSFER: FIRM-LEVEL EVIDENCE FROM CHINESE INDUSTRY

Albert G. Z. Hu, Gary H. Jefferson, and Qian Jinchang*

Abstract—In bridging the technology gap with the OECD nations, developing economies have access to three avenues of technological advance: domestic R&D, technology transfer, and foreign direct investment. This paper examines the contributions of each of these avenues, as well as their interactions, to productivity within Chinese industry. Based on a large data set for China's large and medium-size enterprises, the estimation results show that in-house R&D significantly complements technology transfer—whether of domestic or foreign origin. Foreign direct investment, which we assume is an important channel of proprietary technology transfer, does not facilitate the transfer of market-mediated foreign technology.

Received for publication February 12, 2003. Revision accepted for publication January 26, 2005.

* National University of Singapore, Brandeis University, and National Bureau of Statistics (Beijing), respectively.

We appreciate the assistance of Messrs. Ma Jingkui and Xu Jianyi of China's National Bureau of Statistics. Their collaboration made this research possible. We also appreciate the support of the National Science Foundation (SES-9905259), the U.S. Department of Energy's Biological and Environmental Research Program (DE-FG02-00ER63030), the National University of Singapore Academic Research Fund (R-122-000-058-112), and the Rockefeller Center at Dartmouth College. The very helpful suggestions of Daron Acemoglu (the editor), Adam Jaffe, Tom Rawski, Ren Caifang, Cha Zhimin, and two anonymous referees are greatly appreciated.

I. Introduction

Along with institutional reform and political stability, technological progress is a critical ingredient for sustained economic growth and catch-up. Within the economies exhibiting substantial catch-up, including those of East and Southeast Asia, many technologically lagging firms learned to innovate by first imitating technologies created in developed economies. Imitation may occur through different channels, including market-mediated purchases of technology, technology transfer from multinational corporations to local subsidiaries or joint ventures, or the reverse engineering of products and capital goods. The relative contribution of these channels to technological advance has varied from country to country. Whereas the Philippines and Thailand have been relatively open to foreign direct investment (FDI), Korea has tended to limit FDI and has relied on foreign technology transfer and indigenous R&D.¹ Over time, with the establishment of formal R&D operations, many firms are making

¹ See, for example, Yusuf and Evenett (2002), Jefferson and Zhong (2004), tables 12, 13), and Westphal (1990).