Appendix A

Technical Aspects of the Gravity and Computable General Equilibrium Models

DEAN DeROSA and JOHN P. GILBERT

Gravity Model

This section presents technical aspects of the gravity model that underscore the trade creation projections discussed in chapter 8. It focuses on dataset construction and estimating techniques.

Dataset Construction

As mentioned in chapter 8, the gravity model analysis is based on a dataset that joins elements of two large datasets developed by other researchers. The first of the large datasets is the extensive gravity model dataset developed by Andrew Rose (2004), which covers aggregate bilateral merchandise trade between 178 countries from 1948 to 1999, with gaps and excluding Taiwan and some centrally planned economies. It is compiled from the International Monetary Fund's (IMF) *Directions of Trade Statistics* database. The bilateral trade figures in the Rose dataset are averages of free-on-board export and cost, insurance, and freight import data in US dollars, deflated by the US consumer price index.

The Rose dataset also includes the core explanatory variables discussed in chapter 8 and identified in tables 8A.1 and 8A.2.[1] The core variables are

Dean DeRosa is a visiting fellow at the Peterson Institute for International Economics. John P. Gilbert is associate professor of economics in the Department of Economics, Utah State University, Logan, Utah.

1. The US-Indonesia trade integration and openness dummy variables were compiled for the present study and added to the dataset by the authors.

drawn from several standard sources, including the Central Intelligence Agency's *World Factbook,* IMF's *International Financial Statistics,* Penn World Table, and World Bank's *World Development Indicators.*[2]

To give the Rose gravity model and dataset somewhat greater analytical depth, the Rose dataset was concorded with bilateral merchandise trade data at the 1-digit Standard International Trade Classification (SITC) level, taken from the highly disaggregated bilateral trade dataset compiled by Feenstra et al. (2005). The dataset covers bilateral trade data from 1962 to 2000, organized by four-digit SITC (revision 2) categories. World trade flows are drawn from UN data sources and are based primarily on reporter-country import data, supplemented by reporter-country export data where import data gaps occur and when possible. The Feenstra et al. dataset covers a somewhat smaller number of trading countries than does the Rose dataset, especially for the period between 1984 and 2000, when the dataset contains the bilateral trade of only 72 countries; however, this still accounts for 98 percent of world exports in that period.

For the present study, the Feenstra et al. world trade data were aggregated to the 1-digit SITC level by country pairs and deflated by the US consumer price index (1983 = 100). Real trade flows were transformed to natural logarithmic terms and integrated with the Rose data using a concordance between the UN (Feenstra et al.) country codes and the IMF (Rose) country codes. In the process, all reported Feenstra et al. adjustments to the UN trade data were accounted for except the estimated redistribution of value added in trade between China and Hong Kong, separately reported in Feenstra et al. (2005). Lost in the process, however, were disaggregated trade flows for the former Soviet bloc countries, some less developed countries, and Taiwan, for which no UN (Feenstra et al.) or IMF (Rose) country codes were available. The separate UN country codes for former West Germany and present-day Germany were merged in the Feenstra et al. dataset before integrating the two datasets, thus preserving pre-1991 observations on the bilateral disaggregated trade of West Germany, including trade with Indonesia.

Finally, the explanatory variables in the Rose dataset are augmented by more up-to-date and comprehensive information about 60 bilateral agreements and regional trade agreements (RTAs) involving the European Union and the United States, other high-income countries, and middle-income and low-income countries worldwide, based on the official notifications to the World Trade Organization (WTO) of the membership and start date of the agreements reported in Crawford and Fiorentino (2005). The explanatory variables in the Rose dataset are also augmented by including a variable for the integrity of business practices by pairs of trad-

2. For the CIA's *World Factbook,* see www.cia.gov. For the IMF's *International Financial Statistics,* see ifs.apdi.net. For the Penn World Table, see pwt.econ.upenn.edu. For the World Bank's *World Development Indicators,* see publications.worldbank.org (all Web sites accessed on December 19, 2006).

ing countries. This last explanatory variable ranks countries by the extent of their corrupt business practices, as based on the perceptions of international businessmen and financial journalists compiled and reported annually since 1995 by Transparency International.[3]

Table A.1 describes the regression variables included in the combined Feenstra et al. and Rose datasets, and table A.2 identifies the bilateral agreements and RTAs covered by the three combined RTA variables constructed for the gravity model analysis in chapter 8.

Estimating Techniques

Estimating gravity models using cross-sectional time-series data presents some difficult problems in econometric methodology. Foremost is the problem of accounting for the possible influence of unobservable explanatory variables (Egger 2002, Hsiao 2003). Following Rose (2004), our gravity model analysis relies mainly on the so-called robust standard errors (RSE) variant of the ordinary least squares (OLS) estimation method to yield parameter estimates and standard errors. Both are adjusted for possible heteroskedasticity of the regression error components on a country pair–wise basis combined with the trade commodity groupings. This method is one approach for dealing with unobserved explanatory variables. A more common approach, also carried out as part of the gravity model analysis, uses the so-called random effects (RE) regression model. The RE regression model provides an explicit and formal econometric framework for dealing with unobserved variables but requires that the observed explanatory variables of the model be independent of the unobserved explanatory variables. Otherwise the RE coefficient estimates are biased.[4]

Finally, the analysis utilizes a technique developed by Thomas Plumper and Vera E. Troeger (2004). The technique, which entails several steps, draws heavily on OLS estimating methods and sharply minimizes the correlation of observed and unobserved explanatory variables. However, as the discussion in chapter 8 reveals, the technique produces some extreme and highly implausible estimates for the coefficients of the core explanatory variables.

Computable General Equilibrium Model

This section presents technical aspects of the CGE model that underscores the trade creation projections that we discussed in chapter 8.

3. See Transparency International's corruption index at www.transparency.org (accessed on December 19, 2006).

4. For a general discussion of the RE model and the problem of correlation with observed explanatory variables, see, e.g., Hsiao (2003).

Table A.1 Gravity model regression variables, 1962–99

Regression variable	Description
Dependent variable	Log value of bilateral trade by one-digit SITC, real US dollars
Distance	Log of distance
Joint GDP	Log of product of real GDPs
Joint GDP per capita	Log of product of real GDPs per capita
Common language	Common language dummy
Common border	Land border dummy
Landlocked	Number of countries landlocked (January 2002)
Island	Number of island countries (January 2002)
Land area	Log of product of land areas
Common colonizer	Dummy for common colonizer post-1945
Colony	Dummy for country pairs currently in colonial relationship
Ever a colony	Dummy for country pairs ever in colonial relationship
Common country	Dummy for same nation/perennial colonies
Currency union	Strict currency union dummy
GSP	GSP dummy
EU and US RTAs	Dummy for European Union and 10 RTAs with EU or US hubs
Other HIC RTAs	Dummy for 10 other high-income country RTAs
MIC and LIC RTAs	Dummy for 39 middle-income and low-income country RTAs
Joint TI index	Log of product of TI corruption index
Indonesia-US trade	Dummy for Indonesia-US trade
US openness	Dummy for US trade with all partners
Indonesia openness	Dummy for Indonesian trade with all partners

GSP = generalized system of preferences
HIC = high-income country
LIC = low-income country
MIC = middle-income country
RTA = regional trade agreement
SITC = Standard International Trade Classification
TI = Transparency International

Note: Dependent variable based on bilateral trade flows drawn from the Feenstra et al. dataset. RTAs and Indonesia-US trade and openness variables constructed by the authors.

Sources: Rose (2004); Feenstra et al. (2005); Crawford and Fiorentino (2005); and Transparency International Corruption Perception Index, www.transparency.org.

Overview

As mentioned in chapter 8, our CGE model uses the Global Trade Analysis Project (GTAP) framework, a publicly available and widely adopted model. The GTAP model is a multiregional, multisectoral model that assumes perfect competition and constant returns to scale. Bilateral trade is handled through the Armington assumption, which treats goods from alternative sources as imperfect substitutes. Import-demand functions are

Table A.2 Regional trade agreements, 1962–99

Regional trade agreement	Date of entry into force	Type of agreement	Combined RTA regression variable
EU (Treaty of Rome)	January 1, 1958	Customs union	EU and US RTAs
EFTA (Stockholm Convention)	May 3, 1960	Free trade agreement	Other HIC RTAs
Central American Common Market	October 12, 1961	Customs union	MIC and LIC RTAs
Tripartite Agreement	April 1, 1968	Preferential arrangement	MIC and LIC RTAs
EU-Switzerland and Liechtenstein	January 1, 1973	Free trade agreement	EU and US RTAs
PTN	February 11, 1973	Preferential arrangement	MIC and LIC RTAs
EU-Iceland	April 1, 1973	Free trade agreement	EU and US RTAs
EU-Norway	July 1, 1973	Free trade agreement	EU and US RTAs
Caricom	August 1, 1973	Customs union	MIC and LIC RTAs
Bangkok Agreement	June 17, 1976	Preferential arrangement	MIC and LIC RTAs
EU-Algeria	July 1, 1976	Free trade agreement	EU and US RTAs
PATCRA	February 1, 1977	Free trade agreement	Other HIC RTAs
Sparteca	January 1, 1981	Preferential arrangement	Other HIC RTAs
LAIA	March 18, 1981	Preferential arrangement	MIC and LIC RTAs
Anzcerta	January 1, 1983	Free trade agreement	Other HIC RTAs
Gulf Cooperation Council	July 1, 1983	Preferential arrangement	MIC and LIC RTAs
US-Israel	August 19, 1985	Free trade agreement	EU and US RTAs
Andean Community	May 25, 1988	Preferential arrangement	MIC and LIC RTAs
GSTP	April 19, 1989	Preferential arrangement	MIC and LIC RTAs
Laos-Thailand	June 20, 1991	Preferential arrangement	MIC and LIC RTAs
Mercosur	November 29, 1991	Customs union	MIC and LIC RTAs
AFTA	January 28, 1992	Preferential arrangement	MIC and LIC RTAs
EFTA-Turkey	April 1, 1992	Free trade agreement	Other HIC RTAs
EFTA-Israel	January 1, 1993	Free trade agreement	Other HIC RTAs
Central European FTA	March 1, 1993	Free trade agreement	MIC and LIC RTAs
Kyrgyz Republic-Armenia	October 27, 1995	Free trade agreement	MIC and LIC RTAs
Kyrgyz Republic-Kazakhstan	November 11, 1995	Free trade agreement	MIC and LIC RTAs
SAPTA	December 7, 1995	Preferential arrangement	MIC and LIC RTAs
Armenia-Moldova	December 21, 1995	Free trade agreement	MIC and LIC RTAs
EU-Turkey	January 1, 1996	Customs union	EU and US RTAs
Georgia-Ukraine	June 4, 1996	Free trade agreement	MIC and LIC RTAs
Armenia-Turkmenistan	July 7, 1996	Free trade agreement	MIC and LIC RTAs
Georgia-Azerbaijan	July 10, 1996	Free trade agreement	MIC and LIC RTAs
Kyrgyz Republic-Moldova	November 21, 1996	Free trade agreement	MIC and LIC RTAs
Armenia-Ukraine	December 18, 1996	Free trade agreement	MIC and LIC RTAs
Canada-Israel	January 1, 1997	Free trade agreement	Other HIC RTAs

(table continues next page)

Table A.2 Regional trade agreements, 1962–99 *(continued)*

Regional trade agreement	Date of entry into force	Type of agreement	Combined RTA regression variable
Israel-Turkey	May 1, 1997	Free trade agreement	MIC and LIC RTAs
Canada-Chile	July 5, 1997	Free trade agreement	Other HIC RTAs
Eurasian Economic Community	October 8, 1997	Customs union	MIC and LIC RTAs
Kyrgyz Republic–Ukraine	January 19, 1998	Free trade agreement	MIC and LIC RTAs
Romania-Turkey	February 1, 1998	Free trade agreement	MIC and LIC RTAs
EU-Tunisia	March 1, 1998	Free trade agreement	EU and US RTAs
Kyrgyz Republic–Uzbekistan	March 20, 1998	Free trade agreement	MIC and LIC RTAs
Georgia-Armenia	November 11, 1998	Free trade agreement	MIC and LIC RTAs
Bulgaria-Turkey	January 1, 1999	Free trade agreement	MIC and LIC RTAs
CEMAC	June 24, 1999	Preferential arrangement	MIC and LIC RTAs
Georgia-Kazakhstan	July 16, 1999	Free trade agreement	MIC and LIC RTAs
Chile-Mexico	August 1, 1999	Free trade agreement	MIC and LIC RTAs
Armenia-Russian Federation	March 25, 1993	Free trade agreement	MIC and LIC RTAs
Kyrgyz Republic–Russian Federation	April 24, 1993	Free trade agreement	MIC and LIC RTAs
EU-Romania	May 1, 1993	Free trade agreement	EU and US RTAs
EFTA-Romania	May 1, 1993	Free trade agreement	Other HIC RTAs
EFTA-Bulgaria	July 1, 1993	Free trade agreement	Other HIC RTAs
Melanesian Spearhead Group	July 22, 1993	Preferential arrangement	MIC and LIC RTAs
EU-Bulgaria	December 31, 1993	Free trade agreement	EU and US RTAs
NAFTA	January 1, 1994	Free trade agreement	EU and US RTAs
Georgia-Russian Federation	May 10, 1994	Free trade agreement	MIC and LIC RTAs
Comesa	December 8, 1994	Preferential arrangement	MIC and LIC RTAs
Commonwealth of Independent States	December 30, 1994	Free trade agreement	MIC and LIC RTAs
Romania-Moldova	January 1, 1995	Free trade agreement	MIC and LIC RTAs

Notes: Abbreviated regional trade agreements are the Agreement on Trade and Commercial Relations between Australia and Papua New Guinea (Patcra), ASEAN Free Trade Area (AFTA), Association of Southeast Asian Nations (ASEAN), Australia-New Zealand Closer Economic Relations Trade Agreement (Anzcerta), Caribbean Community and Common Market (Caricom), Common Market for Eastern and Southern Africa (Comesa), Economic and Monetary Community of Central Africa (CEMAC), European Free Trade Association (EFTA), European Union (EU), General System of Trade Preferences among Developing Countries (GSTP), Latin America Integration Association (LAIA), North American Free Trade Agreement (NAFTA), Protocol Relating to Trade Negotiations among Developing Countries (PTN), South Asian Preferential Trade Arrangement (SAPTA), South Pacific Regional Trade and Economic Cooperation (Sparteca), and the Southern Core Common Market (Mercosur).

Sources: Crawford and Fiorentino (2005) and table A.1.

separated by agent, under what is sometimes called the Salter specification.[5] Production conditions are modeled using nested constant elasticity of substitution (CES) functions,[6] and intermediate goods are used in fixed proportions.[7] Representative household demand accounts for changes in demand structures as incomes rise.[8] Recent surveys of the application of CGE models to regional trade negotiations include Scollay and Gilbert (2000), Scollay and Gilbert (2001), Gilbert and Wahl (2002), Robinson and Thierfelder (2002), and Lloyd and MacLaren (2004).

Base Data

The base data for the simulations are drawn from the GTAP6 database (final release), the most complete dataset available. They represent the world economy as it was in 2001. The database contains IO representations of individual economies, obtained from national statistical agencies as well as international trade and income data from the UN Commodity Trade (Comtrade) database and the World Bank, respectively. The GTAP6 database improves significantly on GTAP5 by incorporating new protection data from the Agricultural Market Access Database (AMAD) and Market Access Map (MacMap) databases. The latter feature bilateral tariffs, so RTAs in place in 2001, such as the North American Free Trade Agreement (NAFTA), are fully integrated. Full documentation of the GTAP6 database can be found in Dimaranan and McDougall (2005).

Aggregation Strategy

The GTAP6 database features 87 regions and 57 sectors, but it must be aggregated for computational efficiency. We have used a 28-region and 23-sector aggregation, constructed by ranking the total exports of the United States and

5. In other words, each agent—aggregate household, government, and business firms in each sector—makes its own choice about how much of each intermediate input to source domestically, and how much to import.

6. The CES function treats primary factors of production—capital, skilled and unskilled labor, natural resources, and land—as imperfect substitutes in the production process, with a single elasticity describing substitutability between all factor pairs. Intermediate inputs are used in fixed proportion to output.

7. For any given proportional change in output, intermediate input will grow by the same proportion. The input-output (IO) coefficients are obtained from the IO tables routinely produced by statistical agencies in most economies and constructed for the few regions where the data are unavailable.

8. These and other aspects of the GTAP model are fully documented in Hertel (1997) and the GTAP Web site at www.gtap.org (accessed on December 19, 2006). Changes in demand structure are modeled by so-called nonhomothetic demand functions.

Table A.3 Aggregation scheme for GTAP6 database

Sectoral aggregation	Regional aggregation
Grains	Australia
Other crops	New Zealand
Other agriculture	China
Forestry and fisheries	Hong Kong
Coal, oil, and gas	Japan
Processed rice	Republic of Korea
Sugar	Bangladesh
Other food products	India
Textiles	Sri Lanka
Wearing apparel	Pakistan
Leather products	Chile
Wood products	Rest of South America
Paper products	CAFTA
Chemicals	Western Europe
Minerals and metals	Switzerland
Fabricated metal products	Eastern Europe
Motor vehicles	Morocco
Other transportation equipment	SACU
Electronic equipment	Indonesia
Machinery and equipment	Malaysia
Other manufactures	Philippines
Nontraded services	Singapore
Traded services	Thailand
	Vietnam
	Canada
	United States
	Mexico
	Rest of world

CAFTA = Central American Free Trade Agreement
SACU = Southern African Customs Union

Indonesia and the bilateral exports of the two countries and then using this ranking, along with natural sectoral groupings. We have also included both processed rice and sugar in the sectoral aggregation because of potential sensitivity over those items. A similar approach was followed for regional aggregation, for which care was taken to include current US partners within NAFTA as well as new and prospective US FTA partners, in addition to Indonesia's regional trading partners. New US partners included are Australia, Chile, Morocco, and Singapore; prospective US partners included are countries in the Central American Free Trade Agreement–Dominican Republic, Korea, Malaysia, the Southern African Customs Union, Pakistan, and Thailand. Egypt is also a prospective US FTA partner but is not available in the GTAP6 database. The aggregation is presented in table A.3.

Data Adjustments

The agricultural protection data in GTAP6 are excellent, but services protection data are limited. Dee et al. (2003) estimate barriers to services trade at the aggregate level. In this study, we split services into traded and nontraded categories, following the classification of Dee et al. (2003) and using their estimates of services barriers. The barriers are implemented using several trade policy instruments—import tax equivalents, export tax equivalents, taxes on output, and taxes on domestic capital. The various tax rates were imposed on the GTAP6 dataset prior to the major simulation using the Altertax procedure, which fixes the current account balance and uses parameters such that all key shares in the model remain constant when the new taxes are imposed, while ensuring that the database remains consistent.