How to Evaluate Computable Models of Trade

Alan Fox, Serge Shikher, Marinos Tsigas

United States International Trade Commission[1]

1 Introduction

Trade economists, especially those working in governments, international organizations, and think tanks are often called upon to estimate the effects of various trade policy changes. An example of a trade policy change is a trade agreement. Typically, this analysis has to be performed before a trade policy change is actually implemented and before its effects occur.

The toolbox that economists call upon to answer these questions holds many tools. One of the most commonly used tools is a Computable General Equilibrium (CGE) model. A CGE model incorporates many economic variables such as output, employment, and trade in many industries and countries. It also takes into account the fact that all economic variables are interconnected so that, for example, a country cannot export more than it produces and a slowdown in the auto manufacturing industry has a negative effect on the steel industry.

CGE models appeared in the 1960s and became popular in 1980s and 1990s. There are now many CGE models in existence, including a new class of CGE models called the general equilibrium gravity models. Some features are common to all CGE models, but there are also important differences.

Given the prevalence of CGE models in policy analysis, it is important to know how well these models do in predicting the effects of policy changes. The policy makers that read the estimates would like to know that these estimates are credible. They also want to know the limitations of the models and estimates. In this chapter we will discuss the methods that can be used to evaluate CGE models of trade and present evaluations of several popular models. We will start with a brief overview and history of CGE models.

2 What are the CGE models?

CGE models describe in detail inter-industry linkages in an economy utilizing input-output (I-O) accounts. I-O models of an economy were first introduced by Wassily Leontief in the early 1930s (Leontief, 1966). After publication of Structure of the American Economy in 1941, Leontief continued working on the development of I-O theory. I-O models focus on inter-industry linkages and compute the required resources to satisfy final demand. For example, the I-O accounts can tell us how much of the output of the steel industry is used by the automobile manufacturing industry.

---

[1] E-mails: alan.fox@usitc.gov, serge.shikher@usitc.gov, and marinos.tsigas@usitc.gov. The views expressed in this paper are strictly those of the authors and do not represent the opinions of the U.S. International Trade Commission or of any of its Commissioners.

Leif Johansen extended the I-O model to include economic agents, such as firms and consumers, and market equilibrium conditions (Johansen, 1960). Thus Johansen provided the first CGE model. CGE models presented significant computational challenges since they include multiple industries and countries. Johansen's model was a linear approximation to the underlying economy and could be solved using linear algebra.

CGE models developed further in 1970s. In the United States, John Shoven at Stanford University and John Whalley contributed trade analyses based on CGE models (Shoven & Whalley, 1992). In 1974 they presented a model with two countries, two consumers, and two goods in each country (Shoven & Whalley, 1984).

In Australia, in the 1970s, Peter Dixon and his colleagues in the IMPACT project developed the ORANI model of the Australian economy (Dixon, Parmenter, Ryland, & Sutton, 1977). ORANI is a large scale model, meaning that it includes many industries - 113 in the first version of the model. ORANI-style models were developed by IMPACT staff and others for several economies around the world.[2]

The use of CGE models for policy analysis became more widespread in the 1980s. Many of the CGE models developed in the 1980s were used early in the next decade to estimate the effects of the North American Free Trade Agreement (NAFTA).

In Canada, in 1984, Richard Harris and David Cox analyzed the cost of protection to the Canadian economy and the potential welfare gains that could be realized if Canada were to eliminate its own trade barriers unilaterally or if there were multilateral free trade. For this analysis Harris and Cox developed a CGE trade model which consisted of 20 manufacturing industries that are characterized by scale economies and imperfect competition plus an additional 9 industries, including agriculture, mining, and services, which are modeled competitively and with constant returns to scale (Harris & Cox, 1984).

Turning to multiregional CGE trade models, in 1979 Alan Deardorff and Robert Stern, at the University of Michigan, analyzed changes in tariffs and quantifiable nontariff barriers negotiated in the Tokyo Round with a CGE model of world production and trade (Deardorff & Stern, 1979). We review their model, called the Michigan Model in more detail below. Other models developed in 1980s include WALRAS and RUNS models developed at the OECD (Burniaux et al, 1988; Burniaux and van der Mensbrugghe, 1991).

The 1990s saw the appearance and increased popularity of the Global Trade Analysis Project (GTAP) model, developed at Purdue University by Thomas Hertel, Marinos Tsigas, and others. This model, widely used today for policy analysis, is reviewed in detail below.

A new class of CGE models appeared in the last 10 years. These CGE models share common features with the popular gravity model, which has been very successful empirically in explaining the volume of trade between countries (Anderson and van Wincoop, 2003). Therefore, these new CGE models are often called the general equilibrium gravity models (Larch and Yotov, 2016). Some of these models focus on the variety of firms engaged in production, resulting in a more realistic representation of trade (Eaton and Kortum, 2002). We consider one of these new CGE models in detail below.

---

[2] The Centre of Policy Studies (CoPS), which succeeded IMPACT, has continued work in CGE modeling.In the early 2000's Peter Dixon and Maureen Rimmer developed the MONASH model of Australia (Dixon & Rimmer, 2002).

2.1 Common features and differences between CGE models

CGE models all share basic structural features drawn broadly from economics, and typically motivated by the Arrow-Debreu model (1954), as elaborated by Arrow and Hahn (1971). A minimum set of agents must be modeled within each economy to simulate how trade policy and other factors influence trade patterns. The building blocks of a CGE model can best be understood through the lens of gross domestic product. Measured by expenditures, GDP is often expressed as $Y = C + I + G + (X-M)$. Consumption (C) is typically represented by a representative household. Investment (I) is undertaken by a representative firm for each sector, with the necessary capital supplied by the household. Government consumption (G) is funded through taxes on the various flows of the economy, and provides a public good enjoyed by the representative household. Lastly, exports (X) and imports (M) tie the domestic economy to the rest of the world through the trade of goods and services.

The building blocks of a CGE model include one or more households per region, one or more goods, one or more producers per region, and at least two regions. For example, a simple model might include one household per region, two factors of production (capital and labor) two regions (home and foreign), and two goods. Modern computable trade models feature many more sectors of the economy and regions, with production, consumption and trade modeled in each region and sector of the global economy.

2.1.1 Production

While firms and production are not explicitly part of the expenditure-side measure of GDP, production underpins each of the elements composing GDP. Firms provide the goods and services consumed by households and government, and also provide intermediate inputs to other firms as part of the production process. Firms also serve as the channel for investment, turning savings into productive capital. Lastly, firms produce goods and services demanded by consumers, governments, and producers abroad through exports. The standard approach to modeling the firm uses a representative firm for each good or service in the economy. The firm uses intermediate inputs of goods and services and factors of production such as capital and labor to produce its own product or service that is sold domestically or abroad.[3]

There are several common assumptions made in CGE models to describe production technology of firms. One assumption is that firms use intermediate inputs and factors of production in fixed proportions. Another assumption is that intermediate inputs and factors are substitutable to some degree. Most CGE models assume constant returns to scale, so that doubling all inputs doubles output. Some models assume increasing returns to scale, so that doubling inputs more than doubles output. These models also assume imperfect competition (see below).

The most common models treat firms within a given sector as perfectly competitive. Under such a specification, firms earn no economic profits: the price of a good is just enough to pay for materials, pay the wage bill, and provide investment with a market rate of return. Models may also treat the firm differently, granting firms some degree of market power in a given sector so that the market price is above the marginal cost of production (which is the outcome under perfect competition). Such

---

[3] Some models differentiate between several types of labor and capital. Some models include land and other natural resources as factors of production. Some recent models allow for heterogeneity of firms among one or more dimensions.

frameworks include monopoly, oligopoly, symmetric monopolistic competition (see Dixit and Stiglitz (1977), Krugman (1979, 1980)), and models of firm heterogeneity (Melitz, 2003).

Goods and services may also be differentiated by firm. This approach comports more closely with our intuition regarding the differentiation of goods and services. Firm-level differentiation is also consistent with a number of models of production, including Dixit-Stiglitz/Krugman, Melitz, and Eaton-Kortum.

2.1.2 Consumption and savings

The simplest representation of private consumption uses a representative household, which maximizes utility from consumption of a basket of goods and services available from domestic producers and, thanks to trade, with foreign producers. Consumption is subject to the budget constraint of the household which, in turn, is determined by income derived from the household's factors of production. The household's investment and savings behavior can be very simple (saving a fixed proportion of income) or sophisticated (optimizing an investment portfolio over time subject to a discount rate). The simplest models treat labor supply as fixed, while more complex models might model how household labor supply responds to changes in wages.

Most trade models have a relatively simple approach to modeling savings and investment. Some models treat savings as fixed, while other set savings as a fixed proportion of overall income, that is, with a costant marginal propensity to save. More complex models, such as dynamic models with a time dimension, employ a more sophisticated model of savings behavior over time, taking into account how rates of return on investments are likely to develop over time.

2.1.3 Trade

The motivation for trade may purely be relative price (as in the Ricardian model), or it may be driven by particular characteristics of the good or service. The simplest neoclassical models (such as the Ricardian model) suggest that a country should be only an importer or an exporter of a given good based on its comparative advantage, but not both. Therefore, these models cannot account for the cross-hauling of goods between countries—that is, why the United States not only exports machine tools abroad, but also imports them from a variety of countries.

The Armington (1969) assumption posits that goods are differentiated by country of origin, such that a French good or service is intrinsically different from a German good or service of the same category. Implementing the Armington assumption is relatively simple, with traded goods and services in a given sector aggregated into a composite, with some substitution as to source. By using this approach, a model can account for the two-way trade between countries. With Armington assumption, home bias - the fact that the great majority of domestic spending goes towards domestically produced goods - is explained by consumer preferences.

Several of the recent CGE models use the gravity equation to explain trade. The gravity equation (or the gravity model) focuses on trade costs to explain trade. Trade costs, which are unobservable, are related to various observable country characteristics, such as distance between countries, commonality of language, and shared border. Recent research has shown that trade costs are large and have a significant effect on the volume and pattern of trade (Anderson and van Wincoop, 2003; Shikher, 2013). In the gravity model, home bias is explained by trade costs rather than preferences. Though the gravity and Armington methodologies produce very similar equations describing trade, they describe and estimate key parameters differently, which can produce different results in CGE simulations.

There are other differences in the models that could affect the results. The Michigan model, for example, uses input-output (I-O) tables that were created many years before NAFTA entered into force. Because of the lag in publishing I-O tables, this is frequently the case in trade policy analysis.

The key parameter in CGE models relating to trade is common to the vast majority of all CGE models. This parameter is called trade elasticity (or Armington elasticity) and it measures the responsiveness of trade to changes in prices of imports. When analyzing simulation results of CGE models of trade, it is important to know the values of these parameters, which sometimes vary across industries.

3 The three CGE models considered in this chapter

In the remainder of the chapter we focus on three popular CGE models: the Brown-Deardorff-Stern (BDS) model, the Global Trade Analysis Project (GTAP) model, and the GE gravity model of Eaton and Kortum. We start by describing the key features of these models and then analyze their performance.

3.1 The Brown-Deardorff-Stern model

Alan Deardorff and Robert Stern of the University of Michigan developed one of the earliest computable models of trade, the Michigan Model of World Production and Trade. The model's roots extend back to 1974–1975, when the Bureau of International Labor Affairs of the U.S. Department of Labor asked Deardorff and Stern to evaluate ongoing work on trade policy and employment being conducted as part of the Tokyo Round of negotiations under the GATT. Two proposals were submitted, one relying on econometric estimation of tariff elasticities, and the second calling for the estimation of a disaggregated model of production and trade focused on the industries and countries most central to the Round's negotiation. The Department of Labor funded the econometric approach to the analysis (Leamer, Stern, and Baum, 1977), but the seed had been planted for the computational approach, which Deardorff and Stern decided to pursue on their own. The model continued to be developed and refined over the next decade, with substantial additional features introduced by Drusilla Brown, in particular monopolistic competition in production. This innovation became known as the Brown Deardorff Stern (BDS) model.

The BDS model features 29 industries and 34 countries/regions to describe world production and trade. The industries are based on the broad ISIC industrial classifications 1 through 9, with special attention paid to manufacturing (ISIC 3), which is decomposed into 21 industries. Over time, the number of traded industries increased from only agriculture and manufacturing industries to encompass mining (ISIC 2) and services (ISIC 4-9).

Production is modeled differently by industry. Agriculture is always treated as perfectly competitive with a representative firm, but utility differentiated by country of origin, employing the Armington assumption to explain cross-hauling of trade. Manufacturing and services are generally treated as monopolistically competitive following Dixit-Stiglitz (1977), with production and utility differentiated by firm. When monopolistic competition was first introduced into the model, certain industries were treated as allowing for no entry, while others allowed for free entry. Four were initially treated as perfectly competitive. Over time, all industries were treated as monopolistically competitive with free entry. Mining and the service sectors (ISIC 2 and 4-9) were initially treated as non-traded, but were later classified as traded.

The aggregate level of employment is held constant in each country, with labor free to move between industries. Relative wages between industries are held fixed.

Each country's balance of trade is fixed across the simulation, treating exchange rates as flexible and investment flows between countries as unchanging. Tariff revenues or any rents from nontariff barriers are distributed back to consumers and treated like any other income.

The various incarnations of the BDS model have been used to model the effects of a number of trade policies from the 1970s through the 1990s, including the Tokyo Round of the GATT (Brown and Stern 1986), the U.S.-Canada FTA (Brown and Stern 1989), the NAFTA (Brown, Deardorff and Stern 1992), the Western Hemisphere Free Trade Agreement (WHFTA) (Brown, Deardorff, Hummels and Stern 1994), and the Uruguay Round (Brown, Deardorff, Fox and Stern 1996).

3.2 The GTAP model

The standard GTAP model is closely related to other CGE models of international trade (for reviews see Shoven and Whalley, 1984 and 1992; Francois and Shiells, 1994). In GTAP Data Base version 9, the world is divided into 140 economies. Each economy is specified with demand and production structures for 57 groups of goods and services. Subject to transportation costs, each economy engages in trade for goods and services with all other economies.

Each economy consists of several economic agents. On the final demand side of the model, a utility-maximizing household purchases commodities and saves part of its income, which consists of labor income, capital income, and net tax collections. On the production side of the model, cost-minimizing producers employ factors of production and use intermediate inputs to supply commodities. In the model, intermediate (and final demand) users of commodities are assumed to differentiate a commodity by its region of origin (i.e., the Armington specification is applied).

The model assumes perfect competition, so all firms are price-takers. Regional prices for goods and services are determined by market clearing in domestic and international markets.

Aggregate investment in new capital goods is represented by the output of a "capital goods" sector. For the world as a whole, the sum of household savings is equal to the sum of investment expenditures.

Integrated into this treatment of production, demand, and trade are economic policies which have regional impacts. The GTAP database currently include taxes and/or subsidies on domestic economic interactions and international trade taxes, import duties and export taxes. These policies affect the equilibrium computed by the model and when they change they induce behavioral changes by producers and consumers in all regions.

3.3 The GE gravity model of Eaton and Kortum

The Eaton and Kortum (2002) model of trade belongs to a class of CGE models called general equilibrium (GE) gravity models because the equation that describes trade in the model is a version of the gravity equation.[4]

In the Eaton-Kortum model, industries consist of many producers, each producing its own good. Producers differ in their productivity: some are very productive, others less so. Exporting is done by the more productive producers. Markets are (perfectly) competitive, as in many other CGE models, so producers charge prices equal to costs. The costs in turn consist of costs of production and costs of delivery. Therefore, trade costs play an important role in the Eaton-Kortum model, just as they do in the gravity model.

Consumers buy products from the producers that offer the best price. As in most CGE models, consumer income consists of labor income and capital income. Model equations insure that supply is equal to demand in product and labor markets.

In the Eaton-Kortum model, the key parameters determining the pattern of trade are trade costs and productivities of producers. Productivities determine comparative advantages, which in turn determine the pattern of trade.

The original Eaton-Kortum model included only one industry. The model we evaluate in this chapter is an extension of the Eaton-Kortum model to multiple industries by Shikher (2011, 2012).

Since each industry consists of many goods, and consumers buy each good from only one producer at a time, there is naturally a two-way trade between countries. For example, Germany buys U.S. cars and the United States buys German cars, but they are not the same cars. Therefore, there is no need for the Armington assumption, which is typical in other CGE models, that says that consumers differentiate products only on the basis of their national origin.

4 Evaluation of computable models

Kydland (1992) listed the following steps in conducting computational experiments.[5] The first step is to formulate a question that needs to be answered by computational experiments. The second step is to choose a model to answer this question. The third step is to assign values to model parameters. This can be done by taking values from other studies or by estimating them using all or part of the model.

The fourth step is to perform robustness checks by varying parameter values within a reasonable range. The fifth and last step is to perform a quantitative assessment of the precision with which the model can answer the question that is being asked. To perform this step, it is necessary to select a set of key facts to which model results can be compared, such as a collection of sample statistics or specific events that one wants the model to replicate. It is also necessary to select a metric for comparing model outcomes with data. The rest of this section will focus on the specifics and peculiarities of performing this assessment for computable models of trade.

---

[4] See Head and Maier (2014) for a good review of the gravity equation literature.

[5] Kydland was a co-recipient of the 2004 Nobel Memorial Prize in Economics.

Kehoe (2005) stresses the need to evaluate applied GE models of trade by matching the results from the models to data. In this chapter, we are guided by two main ideas: (a) predictions of trade models need to be evaluated ex-post and (b) trade models need to be able to accurately predict at least some aspects of changes in trade that occur because of policy changes.

4.1 In-sample fit vs. out-of-sample forecasting

The first thing to note about computable trade models is that they typically have many parameters, including preferences, technology, trade costs, and others. In fact, there are oftentimes as many parameters as data points used to estimate the parameters. Such models have a perfect or nearly perfect fit with the data used to parameterize them ("in-sample fit"). Therefore, evaluating in-sample fit of models in this case is not indicative of the quality of these models. Instead, it is useful to take the models outside the time interval used to parameterize them. In other words, data used to estimate parameter values should be different from the data used to evaluate the model. By doing this, we can learn which aspects of the data the model can replicate (Canova & Ortega, 2000; Kehoe, 2005).

In addition, with multi-equation models, statistical fit of each equation and the ability of the whole model to track the data do not necessarily go hand-in-hand. For example, each equation may have a good fit while the model as a whole may estimate data poorly (Pindyck and Rubinfeld). What we are interested in, of course, is the ability of the whole model to produce accurate estimates.

Evaluation of models by out-of-sample data is common in another field of economics, Real Business Cycle literature, which followed Kydland's suggested steps above, but is extremely rare in the trade literature.

There are several approaches that can be used to evaluate out-of-sample model performance. One is to make forecasts about the future and wait to see if they turn out to be correct. While this is a reasonable approach in finance or weather forecasting where forecast periods are short, it is not practical in international trade where forecast periods are measured in years or decades. More practical approaches are historical forecasting and backcasting, covered in the next section.

4.2 Historical forecasting and backcasting

One practical approach to evaluating out-of-sample performance of computable models of international trade is called historical forecasting. With this approach, a model is parameterized using data up to a certain point in the past. Let's call this point T_0. The model is then used to estimate variables of interest at some point, T_1, which is still in the past but occurs after T_0.

Typically, T_0 is chosen because some important event "X" affecting trade occurred on that date. T_1 is chosen because the effects of X are expected to be fully felt by then. As a robustness check, several different T_1 dates can be tried.[6]

For example, if we want to see how well a model can estimate effects of trade agreements, we can check how well it could estimate the effects of NAFTA. In that case, T_0 is January of 1994, when NAFTA

---

[6] Choosing T_1 that is further from T_0 than needed would make the analysis less powerful since it would make it more likely that other events, not included in the model, affect international trade.

came into force and T_1 can be, for example, end of year 2000, by which time the vast majority of NAFTA's provision had been implemented.

With backcasting, the model is parameterized using current data and then asked to predict the past. In other words, T_0 is in the present while T_1 is in the past. Backcasting is easier to implement because one can just take a current model and use it without reparametrizing it with historical data, as required by historical forecasting. An example of backcasting would be using a CGE model parameterized with current data to predict the economy of 1990. Model predictions can then be compared to observed data to evaluate model performance.

4.3 Which moments of the data should trade models be able to replicate?

One of the key questions a trade modeler must answer is which aspects of the trade data should the model be targeted to replicate. Since trade models are designed to study international trade, we would like them to be good at estimating the effects of various policy changes on international trade. However, even within the realm of international trade there are many variables to look at. Two basic aspects of trade are the volume of trade (how much is traded) and pattern of trade (who sells what to whom).

When measuring the volume of trade, one obvious variable is the volume of trade between two countries. However, the volume of trade is sensitive to total economic activity in each country. For example, trade declines in recessions and increases in economic expansions. To correct for this effect, we can measure trade relative to GDP or total spending in the importing country.

A trade model should also be able to correctly estimate the pattern of trade. We suggest using import shares to measure the pattern of trade. Country i's share of country *n*'s imports of good *j*, $IMPSHARE_{inj}$, is equal to the value of country n's imports of *j* from *i*, $IMP_{inj,}$, divided by total imports of *j* by *n*, IMP_{nj}: $IMPSHARE_{inj}=IMP_{inj}/IMP_{nj}$.

Import shares can be affected by all the things that affect the pattern of trade. Changes in comparative advantages would affect the pattern of trade. So would changes in relative trade costs.[7] However, comparative advantages change very slowly and most determinants of trade costs, such as distance or commonality of language, are also fairly stable over time. Therefore, major changes in trade policy that affect trade costs are strong candidates for explaining changes in import shares in the short and also medium term.

An alternative is to measure trade relative to GDP, as done in Kehoe (2005). Trade relative to GDP is a good measure of the volume of trade while import shares are better suited for measuring the pattern of trade. Trade to GDP ratios can also be affected by recessions and greater trade in intermediate goods.[8]

We also need a metric for comparing model predictions with data. Following Kehoe (2005) we regress the actual changes in a variable on its predicted changes. Ideally, the intercept of such a regression would be 0, while the slope and R^2 would equal 1. Deviations from these ideal values would tell us about the performance of the model.

---

[7] If trade costs change for all countries and industries, for example because shipping becomes cheaper, then import shares would not be significantly affected.

[8] As production spreads over more countries, total volume of trade increases as intermediate goods cross international borders multiple times before becoming final products.

4.4 Ex-post decomposition vs. historical forecasting

International trade modelers use historical data in a number of ways. We have reviewed historical forecasting and backcasting, which can be used to evaluate model performance. Alternatively, historical data together with a model can be used to analyze historical changes in the data. When CGE models are used for this purpose, it is called ex-post decomposition. In this procedure, model parameters are picked so that the model fits the data between T_0 and T_1 perfectly.[9] The parameters are often allowed to vary between T_0 and T_1 in this exercise.

The key difference between ex-post decomposition and historical forecasting is that in ex-post decomposition, all historical data is used to parameterize the model while in historical forecasting only the data up to a certain time in the past is used to parameterize the model and the rest of the data is used to evaluate model predictions. In ex-post decomposition, changes in key parameters and variables of the model that occur between T_0 and T_1 provide an explanation of how the economy arrived from T_0 to T_1. The results of ex-post decomposition have to be interpreted with caution, however. Explanations of historical changes provided by this exercise are only right if the model is correct.

5 NAFTA as a historical experiment

We use the North American Free Trade Agreement (NAFTA) as a large-scale historical experiment to compare the performance of the three models of trade described in Section 3. We parameterize all three models using pre-NAFTA data and simulate tariff removal scheduled under NAFTA. We then compare model predictions with post-NAFTA observations.

Possible effects of NAFTA were studied before NAFTA by several groups of researchers.[10] CGE models were heavily utilized in making predictions regarding the effects of NAFTA. The models utilized the Armington assumption and assumed either constant or increasing returns to scale.[11] Some models had constant capital stock, while others allowed capital accumulation. Some models only simulated removal of tariffs while others also simulated removal of non-tariff barriers.[12]

Post-NAFTA studies of the effects of NAFTA have typically utilized the gravity equation rather than CGE models. [13] These studies generally found that NAFTA had a relatively small effect on employment, prices, and welfare, as pre-NAFTA studies predicted. They also found that NAFTA had a large effect on trade, which is different from the pre-NAFTA predictions made by CGE models.

---

[9] Remember that CGE models can fit historical data perfectly or nearly perfectly since they have many parameters.

[10] The pre-NAFTA studies were initially collected together by the USITC (1992). The updated versions of some of these studies, together with the several new ones were later collected in Francois and Shiells (1994).

[11] See Baldwin and Venables (1995) for a review and classification of these models.

[12] Assuming increasing returns to scale instead of constant returns yielded greater predicted effects of NAFTA. Allowing international movement of capital typically caused large inflows of capital into Mexico. Removing non-tariff barriers in addition to tariffs results in greater predicted effects of NAFTA.

[13] The examples include Gould (1998) and Krueger (1999). Unfortunately, many of these studies do not use the theoretically-derived specification of the gravity equation of Anderson and van Wincoop (2003). Romalis (2007) uses a CGE model with the Armington assumption, parameterized with the post-NAFTA data, to study the effects of NAFTA. Reviews of the post-NAFTA literature can be found in Burfisher et al (2001) and Romalis (2007).

With few exceptions, pre-NAFTA predictions made by economic models have not been evaluated. One paper that evaluates the performance of pre-NAFTA predictions is Kehoe (2005).[14] By systematically comparing model predictions to data, he finds that many of the predictions made before NAFTA turned out to be significantly off.[15] Specifically, the pre-NAFTA forecasts significantly underestimated the effects of NAFTA on trade, sometimes by several orders of magnitude. In addition, the models did poorly in explaining the variation of changes in trade flows across countries and industries. Analysis performed in Shikher (2012) supports these conclusions.

5.1 Basic facts about NAFTA

The North American Free Trade Agreement was signed by Presidents George H.W. Bush and Carlos Salinas de Gortari, and Prime Minister Brian Mulroney on December 17, 1992. The U.S. Congress passed implementing legislation in November 1993. President Bill Clinton signed the NAFTA bill on December 8, 1993 and the agreement entered into force on January 1, 1994.

Prior to NAFTA, Canada and the United States had negotiated the Canada-U.S. Free Trade Agreement (CUSFTA) which entered into force on January 1, 1989. Many bilateral tariffs were removed immediately when CUSFTA took effect; the agreement was fully implemented by 1998.

NAFTA gave Mexican products duty-free access to the U.S. and Canadian economies, and opened the Mexican economy to imports from the United States and Canada. Many tariffs were removed immediately, while other tariffs were gradually removed over 15 years.

NAFTA and CUSFTA represented the latest in a series of initiatives to integrate North American markets. In 1911 President Taft and Prime Minister Laurier signed a trade agreement between the United States and Canada. This agreement was short-lived, however, as a new Canadian government rejected it in late 1911. In 1965, the United States and Canada liberalized bilateral trade in motor vehicles and parts under the Canada–United States Automotive Products Agreement, or Auto Pact.

Mexico had been reforming its economic policies since the 1980s, abandoning import substitution policies and joining the General Agreement on Tariffs and Trade (GATT) in 1986. For the United States, a trade agreement with Mexico was expected to expand exports to a growing market of almost 100 million people. For Mexico, a trade agreement with the United States was expected to contribute to consolidating earlier market opening measures and to attract greater flows of foreign investment in Mexico. It was also thought that NAFTA could incentivize conclusion of the Uruguay Round of multilateral trade negotiations which had started in 1986 and concluded in 1994.

5.2 Evaluating model performance

All three models were parameterized with pre-NAFTA data. The latest version of the GTAP model was parameterized with version 3 of the GTAP database, which is based on 1992 data. That database has 8 manufacturing industries as well as several nonmanufacturing industries. It has 5 "countries": Canada, Mexico, USA, EU12, and the rest of the world (ROW). The multi-industry EK model was parameterized with 1989 data (Shikher, 2012). It has 8 manufacturing industries and 19 OECD countries. The industrial classifications in the EK and GTAP models are the same. The Michigan model was parameterized with

---

[14] Fox (2000) evaluates the performance of the BDS model in predicting the effects of the CUSFTA.

[15] Kehoe reviews the forecasts of the Brown-Deardorff-Stern, Cox-Harris and Sobarzo models.

1989 data. It has 29 industries (of which 21 are in manufacturing) and 5 countries: U.S., Canada, Mexico, "31 other major trading countries", and the ROW.

With all models, the simulations entailed removing tariffs and tariff equivalents of non-tariff barriers (NTBs) between NAFTA countries.[16] Only trade in manufacturing industries is liberalized; other industries are left as is. The same amount of protection is removed in GTAP and EK models. Tariffs and NTB tariff equivalents used in these two models are from Nicita and Olarreaga's Trade, Production, and Protection Database (2006). The Michigan model has its own estimates of protection that predate the NAFTA agreement.[17]

We performed 4 versions of GTAP simulations: (a) using standard GTAP industry-specific trade elasticities (which vary between 2 and 7), (b) with all trade elasticities set to 8 (both domestic/import and import sources substitution elasticities), (c) with all elasticities set to 3 (to make it similar to the BDS model), and (d) with standard GTAP elasticities and Johansen's solution method, which is a linear approximation method that was common in 1980s and may produce less accurate results.

To evaluate predictions, we look at (a) changes in import shares (shares of country *i* in country *n*'s imports of industry *j*) and (b) changes in NAFTA trade (relative to total trade or GDP). We will compare model predictions to the actual changes during 1989-2008. In order to make the comparisons between the predicted and actual changes, we will use the following comparison metrics following Kehoe (2005): correlation between predicted and actual values, and intercept and slope from the regression of actual on predicted values.[18]

Table 1 Summary of results

Measure	Actual 1989-2008	Estimated EK	Estimated GTAP(std)
NAFTA trade relative to the total trade of the NAFTA countries	24.8%	25.9%	28.7%
NAFTA trade relative to the total income of the NAFTA countries	66.5%	62.2%	48.0%

Note: NAFTA trade is the sum of all bilateral trade flows in manufacturing goods between the NAFTA countries. The total trade of the NAFTA countries is the sum of their manufacturing exports and imports with all countries. The total income of the NAFTA countries is the sum of their GDPs. EK simulation results are from Shikher (2012). GTAP results are produced with standard GTAP trade elasticities.

We realize that model predictions can deviate from actual changes because events other than NAFTA occurred between 1989 and 2008. We will discuss these events in Section 5.3.

The first row of Table 1 shows changes in the share of NAFTA trade in the total trade of the NAFTA countries. This share increased between 1989 and 2008 by 24.8%. The EK model predicts that it would grow 25.9%, which is very close, while the standard GTAP model predicts it to grow 28.7%, which is a little higher, but still close. The second row shows NAFTA trade relative to the total income of the NAFTA countries. It increased 66.5% between 1989 and 2008 while the EK model predicts it to increase 62.2%. The prediction of the GTAP model with standard elasticities is lower at 48%.

---

[16] Magnitudes of tariffs and tariff equivalents of non-tariff barriers are shown in Table 6 in the appendix.

[17] All results for the BDS model published here are derived from Brown, Deardorff and Stern (1992). Because certain data are not featured in the article, some comparisons with the EK and GTAP models cannot be derived.

[18] The R^2 for this regression is correlation squared.

Table 2 Estimated vs. actual percent changes in trade

Variable	Actual 1989–2008	Estimated EK	Estimated GTAP(std)	Estimated BDS
Canadian exports	66.7	45.4	23.6	4.3
Canadian imports	58.2	37.1	16.9	4.2
Mexican exports	120.3	130.4	79.4	50.8
Mexican imports	64.2	58.3	42.2	34.0
U.S. exports	39.2	24.0	11.0	2.9
U.S. imports	46.2	17.5	7.7	2.3
Correlation with data		0.98	0.95	0.86

Note: Exports and imports are total exports and imports of manufacturing goods relative to GDP. EK results are from Shikher (2012). GTAP results are produced with standard GTAP trade elasticities.

Table 2 shows changes in the NAFTA countries' total exports and imports due to NAFTA. The first column of numbers shows that exports and imports grew in all NAFTA countries, most of all in Mexico, where exports more than doubled between 1989 and 2008. The other columns of numbers show the estimates of the three models: EK, GTAP, and BDS. We can see that BDS has generally underpredicted changed in exports and imports of the NAFTA countries, which is common to pre-NAFTA predictions. Compared to BDS's, GTAP's estimates are closer in magnitude to those observed in the data while the estimates of the EK model are the closest. The last row shows the correlation between the actual and predicted total exports and imports for each model. It shows how well each model can explain the variation in the total trade across countries. The EK model has the highest correction with GTAP a close second.

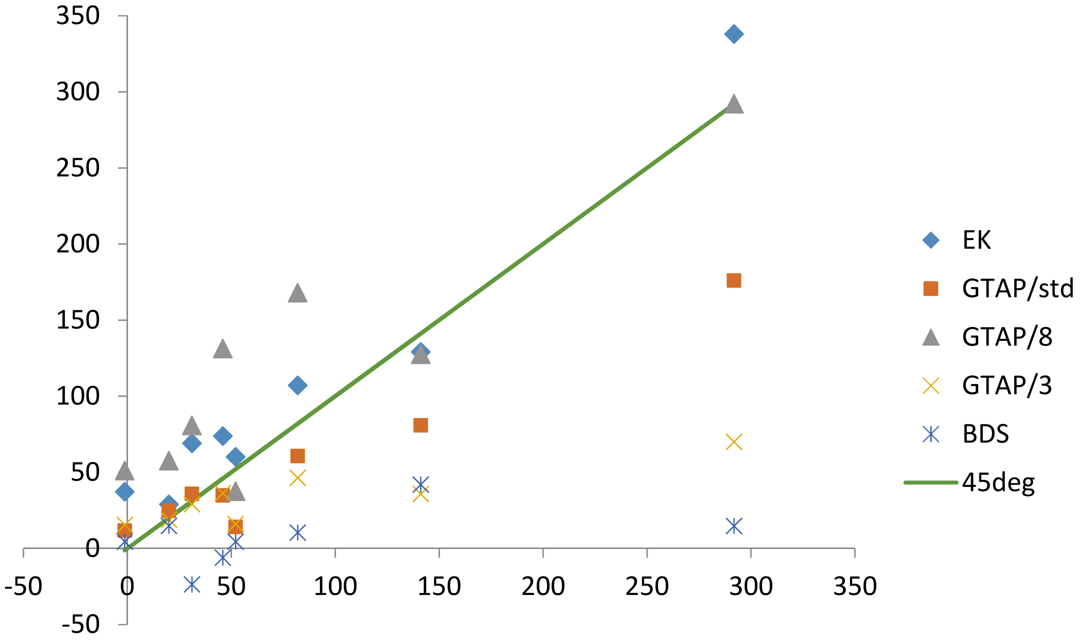
Table 3 Actual percent changes in import shares, 1989-2008

Importer	Exporter	Food	Textile	Wood	Paper	Chemicals	Nonmetals	Metals	Machinery
Canada	Mexico	92.7	202.8	1580.5	185.8	413.4	208.9	-59.7	254.1
Canada	U.S.	36.1	72.8	16.1	9.0	14.6	26.0	6.8	2.8
Mexico	Canada	20.8	-65.0	846.4	-40.8	5.4	-68.9	-70.2	-35.6
Mexico	U.S.	18.3	22.9	-11.0	-1.4	8.6	3.9	-7.8	4.7
U.S.	Canada	73.4	93.7	-4.4	-11.2	5.8	17.7	-9.1	-5.7
U.S.	Mexico	81.9	291.8	52.1	-1.3	45.9	31.0	20.0	141.1

Next, we investigate the accuracy of the models' forecasts at the industry level. Table 3 shows the actual percentage changes in the import shares for each pair of NAFTA countries by industry. The numbers in the table are calculated as trade divided by total imports of the importer. Changes predicted by the models are shown in Table 7 in the appendix.

Figure 1 graphs the predicted vs. actual Mexican shares of U.S. imports. The actual changes are on the horizontal axis while the predicted changes are on the vertical axis. We also plot the 45-degree line to show the ideal location for the points on the graph. We can see that the EK model overpredicts changes in import shares, but generally stays close to the 45-degree line. The GTAP model with standard GTAP elasticities matches changes in import shares well when they are small, but underpredicts changes when changes are medium or large. The GTAP model with elasticities set to 3 also matches small changes well, but significantly underpredicts large changes in trade. With all elasticities set to 8, the GTAP model overpredicts small changes but matches well the large changes. The Michigan (BDS) model performs poorly and tends to significantly underpredict changes in trade.

Figure 1 Estimated vs. actual percent changes in the shares of Mexican goods in U.S. imports (each point represents an industry)



GTAP simulations use the following elasticites: GTAP/std uses standard GTAP elasticities, GTAP/8 uses all elasticities equal to 8, GTAP/3 uses all elasticities equal to 3.

Table 4 and Table 5 summarize how well the predicted changes in industry-specific trade shares match the actual. Table 4 looks at the variation of trade flows across industries for each importer-exporter pair. As explained in the previous section, to compare the actual and predicted changes, we regress the actual changes on predicted. Ideally, the coefficient of this regression would be 1, slope 0, and correlation (which is the square root of R^2) 1. We can see that the EK model does a good job predicting trade changes, except for trade between Canada and Mexico. In fact, all of the models produce large errors when predicting trade changes between those countries, which is a puzzle. It should be noted the Canada-Mexico trade is about 1% of total NAFTA trade. The pre-NAFTA trade between those countries was very small and post-NAFTA changes, in percentage terms, were substantial. The large prediction errors might be due to measurement errors in the data, which are more likely for small trade flows, or due to poor ability of the models to predict trade changes starting from very low initial levels of trade.[19]

The EK model does well predicting changes in trade between export-importer pairs other than Canada-Mexico. The slope is close to one and correlation is between 0.72 and 0.98. The BDS model underpredicts changes in the U.S. imports from Mexico and Canadian imports from the U.S. (slope is much greater than 1). At the same time, it overpredicts changes in the U.S. exports to Mexico and U.S.

---

[19] Canadian exports to Mexico were only $400M in 1989 and $1B in 2000. In 1989, Mexican Wood Products exports to Canada were reported to be only $8.2M. Low initial trade results in very large percentage changes post-NAFTA: exports of Wood Products from Mexico to Canada grew by 1385%.

imports from Canada (slope is much less than 1). The correlation across industries between actual and predicted trade changes is generally low.

Table 4 Comparisons of actual (1989-2008) and estimated changes in import shares

		EK model			BDS model		
Importer	Exporter	Correlation	Intercept	Slope	Correlation	Intercept	Slope
Canada	Mexico	-0.15	423.10	-1.31	0.41	111.09	23.89
Canada	U.S.	0.91	5.71	1.04	0.95	5.54	2.88
Mexico	Canada	-0.57	-185.64	-12.53	-0.14	93.82	-0.81
Mexico	U.S.	0.72	-9.46	1.00	0.10	2.54	0.31
U.S.	Canada	0.77	-7.59	0.81	0.28	12.26	0.58
U.S.	Mexico	0.98	-15.70	0.93	0.44	65.84	2.23

		GTAP(std) model			GTAP(8) model		
Importer	Exporter	Correlation	Intercept	Slope	Correlation	Intercept	Slope
Canada	Mexico	-0.23	458.71	-2.51	-0.29	558.96	-2.28
Canada	U.S.	0.90	9.69	0.51	0.92	7.47	0.47
Mexico	Canada	-0.49	295.40	-13.58	-0.47	336.41	-10.99
Mexico	U.S.	0.78	-13.37	0.79	0.79	-17.41	0.80
U.S.	Canada	0.86	-4.69	0.76	0.91	-10.79	0.60
U.S.	Mexico	0.98	-11.24	1.72	0.90	-38.32	1.02

		GTAP(3) model			GTAP(std,Joh) model		
Importer	Exporter	Correlation	Intercept	Slope	Correlation	Intercept	Slope
Canada	Mexico	-0.20	520.72	-5.65	-0.22	473.98	-4.44
Canada	U.S.	0.92	5.96	1.27	0.90	7.12	0.81
Mexico	Canada	-0.46	351.60	-18.57	-0.45	284.69	-11.31
Mexico	U.S.	0.85	-15.92	1.31	0.75	-11.90	0.73
U.S.	Canada	0.93	-19.25	2.29	0.88	-12.13	1.40
U.S.	Mexico	0.89	-69.73	4.56	0.97	-37.87	3.46

Note: GTAP(std) uses standard GTAP elasticities, GTAP(8) uses all elasticities equal to 8, GTAP(3) uses all elasticities equal to 3, GTAP(std,Joh) uses standard GTAP elasticities and Johansen's solution method.

Next, we look at the performance of the GTAP model. When used with its standard elasticities, the GTAP model tends to underpredict changes in the U.S. imports from Mexico and overpredict all other trade changes. The correlation across industries for the four pairs of countries (excluding Canada-Mexico trade) is high.

GTAP performs better when all elasticities are set to equal to 8. The slope is 1 for changes in the U.S. imports from Mexico. However, the model overpredicts changes in the U.S. imports from Canada. The model also still overpredicts changes in the Canadian imports from the U.S. When all elasticities are set to 3, the GTAP significantly underpredicts all trade changes. The cross-industry correlations are high for all country pairs, except Canada-Mexico trade.

Interestingly, the GTAP model with standard elasticities performs noticeably worse when solved using the Johansen's linear approximation method. While non-linear solution methods are widespread now, the linear solution methods were popular in the early days of the CGE modeling because they require significantly less computing power. Our results suggest that errors introduced by linear approximation methods can be significant. The BDS model was also solved using Johansen’s method.

Table 5 Correlations between actual and estimated industry-level bilateral trade flows (32 observations, excluding trade flows between Canada and Mexico)

Multi-industry Eaton-Kortum	0.95
GTAP with standard elasticities	0.86
GTAP with elasticities equal to 8	0.90
GTAP with elasticities equal to 3	0.83
GTAP with standard elasticities and Johansen's solution method	0.79
Brown-Deardorff-Stern	0.31

Table 5 shows correlations between predicted and actual trade changes for all models. The correlations are calculated for all importer-exporter pairs except Canada-Mexico. This results in 32 trade flows (4 country pairs with 8 industries each). The EK model does the best with correlation of 0.95. The GTAP model with elasticities set to 8 also does well, with correlation equal to 0.9. The correlation declines to 0.86 with GTAP's standard elasticities are used and to 0.83 when all elasticities are set to 3. The correlation declines further to 0.79 when Johansen's solution method is used. The BDS model has the lowest correlation, 0.31.

5.3 Analysis of performance differences

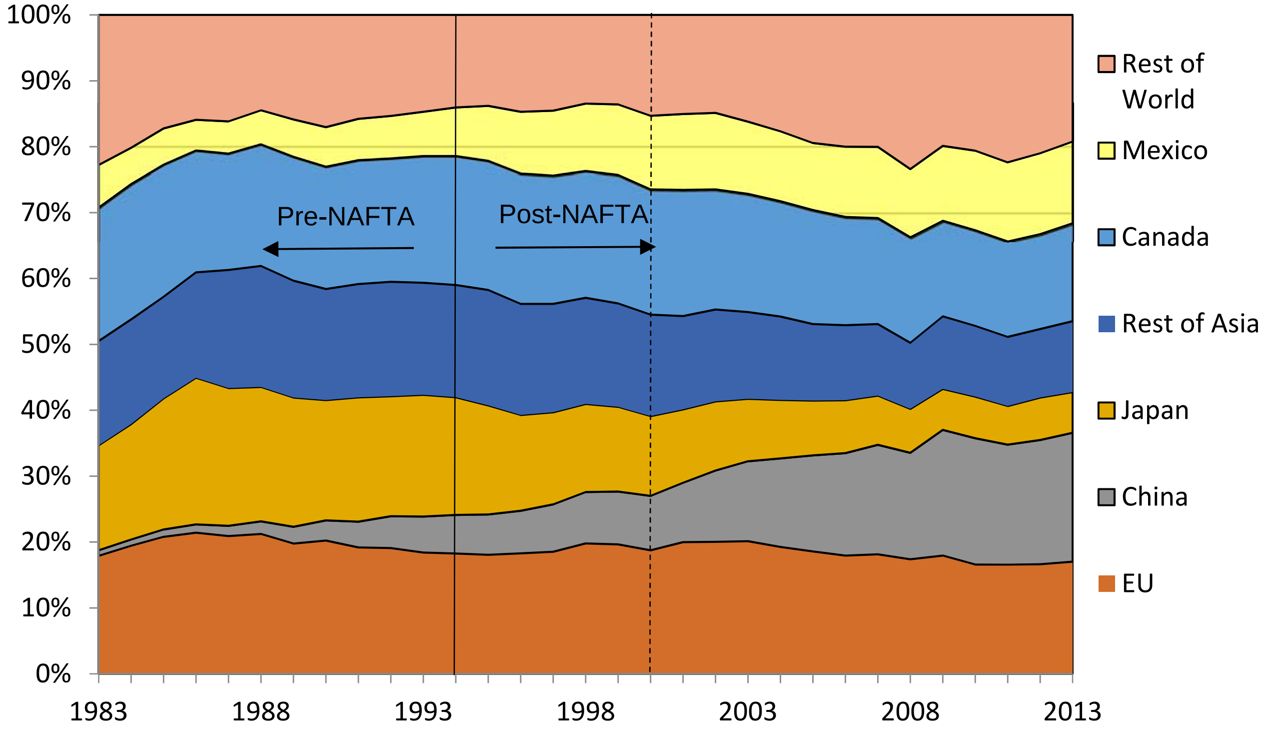
What causes the differences in performance between the models? The differences can come from model assumptions or parameter values. For example, we saw that the performance of the GTAP model improves when its trade elasticities are changed from their standard values to 8, as in the Eaton-Kortum model. In some instances, parameters may be obtained from old data. For example, input-output tables used in CGE models are difficult to create and therefore rarely updated. Sometimes, they are decades old. Models also differ in some assumptions, but sometimes different assumptions produce virtually identical model equations.[20]

The predictions of the models can deviate from the data because the events that are outside the scope of the models occur between T_0 and T_1. For example, there could be technological changes, which are typically not modeled by CGE models of trade. However, the pattern of trade is determined by comparative advantages, which are productivities measured relative to other industries and other countries. For example, the pattern of trade is not affected when productivities in all industries and countries increase equally. Changes in comparative advantages occur slowly, so they are unlikely to explain large increases in NAFTA trade.

---

[20] As mentioned before, this is the case with the Armington assumption and Eaton-Kortum methodology.

Figure 2 U.S. Merchandise Imports by Country and Region, 1983-2013

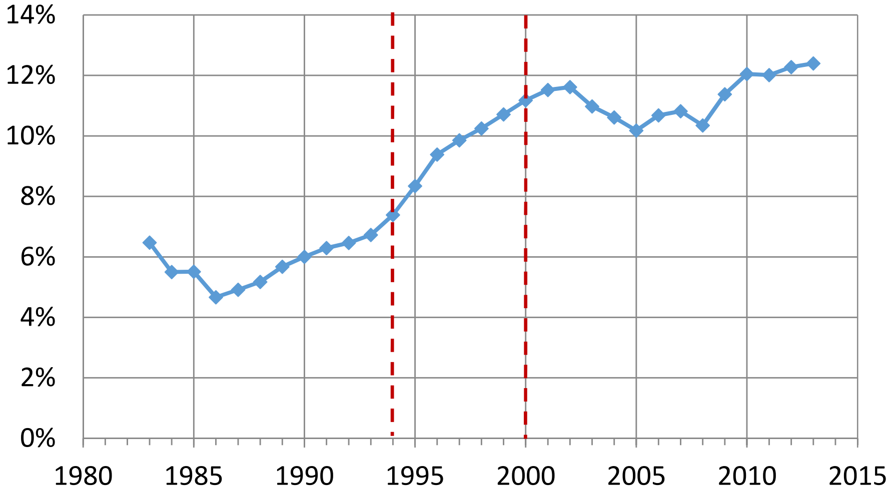


Source: De La Cruz and Riker (2014) using data from the U.S. Department of Commerce.

Several events happened between the implementation of NAFTA in 1994 and 2008. Mexican Peso devaluation occurred less than 12 months after NAFTA entered into force. Everything else equal, peso devaluation increases Mexican exports and decreases Mexican imports. There are several reasons why we think that devaluation does not have a large effect on the analysis above. The devaluation affected all industries, so it should not have affected cross-industry variation of Mexican shares in U.S. imports. More importantly, we expect that the effects of the devaluation had dissipated by 2008.

Another major phenomenon that occurred between 1989 and 2008 is the rise of China. There are several reasons why we think that our results should not be significantly affected by this phenomenon. As shown in Figure 2, Chinese import share in the U.S. increased mostly at the expense of Japanese and other Asian countries' import shares. The Mexican share in U.S. imports grew between 1994 and 2000 (Figure 3) and then remained constant even though Chinese share continued to increase. The share has not changed much between 1983, when it was 6.5%, and 1993, when it was 6.7%. Then it grew to 11.2% in 2000 and remained close to that level in 2008, when it was 10.3%. Most provisions of NAFTA have been implemented by the year 2000.

Figure 3 Mexico's share in U.S. imports, in percent



Source: De La Cruz and Riker (2014) using data from the U.S. Department of Commerce.

6 Discussion of the results

Computable general equilibrium models are an important tool for policy analysis. Their importance will most likely increase further in the years to come as computing power of hardware and user friendliness of software continue to improve. At the same time, new trade deals and import restraints continue to appear every day, which fuels the demand for policy analysis that CGE models satisfy.

CGE models are different from many other economic models due to their complexity and high number of parameters. Critics often say that CGE models are "black boxes" meaning that it is difficult to get an intuition for the results they produce. However, this complexity is sometimes unavoidable when modeling the behavior of the whole economy.

The complexity and high number of parameters of CGE models dictate that they need to be evaluated (a) in their entirety rather than individual equations and (b) using data that was not used to parameterize the models. We suggest that historical forecasts are a good tool for model evaluation. It is commonly used in other fields of economics and other sciences and we hope that its use would increase in international trade modeling.

In this chapter, we describe the nature of CGE models and their history. We then focus on three important CGE models of trade: the Brown-Deardorff-Stern model, the GTAP model, and the general equilibrium gravity model of Eaton and Kortum extended to multiple industries. We used the models to perform historical forecasts of NAFTA and then compared model predictions with post-NAFTA data.

We have discussed some reasons for differences in model forecasts, such as trade elasticities. More work is needed to evaluate CGE models of trade and understand how different model elements contribute to different model predictions.

7 Bibliography

Anderson, J. E. (1979). A Theoretical Foundation for the Gravity Equation. *American Economic Review , 69* (1), 106-16.

Anderson, J. E., & van Wincoop, E. (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review , 93* (1).

Armington, P. S. (1969). A Theory of Demand for Products Distinguished by Place of Production. *IMF Staff Papers , 16* (1), 159-78.

Armington, P. S. (1969). The geographic pattern of trade and the effects of price effects. *IMF Staff Papers* .

Arrow, K. J., & Debreu, G. (1954). Existence of an Equilibrium for a Competitive Economy. *Econometrica , 22* (3), 265-290.

Arrow, K., & Hahn, F. (1971). *General competitive analysis.* Holden-Day.

Baldwin, R. E., & Venables, A. J. (1995). Regional Economic Integration. In G. M. Grossman, & K. Rogoff (Eds.), *Handbook of International Economics, Volume III.* Elsevier.

Bergstrand, J. H. (1985). The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence. *Review of Economics and Statistics , 67*.

Brown, D. K., & Stern, R. M. (1989). Computable general equilibrium estimates of the gains from U.S.-Canadian Trade liberalization. In D. Greenaway, T. Hyclak, & R. J. Thornton (Eds.), *Economic Aspects of Regional Trading Arrangements* (pp. 69-108). Hertfordshire, England: Harvester Wheatsheaf.

Brown, D. K., Deardorff, A. V., & Stern, R. M. (1992). A North American Free Trade Agreement: Analytical Issues and a Computational Assessment. *The World Economy , 15* (1), 15-29.

Brown, D. K., Deardorff, A. V., Fox, A. K., & Stern, R. M. (1996). The liberalization of services trade: potential impacts in the aftermath of the Uruguay Round. In W. Martin, & L. A. Winters (Eds.), *The Uruguay Round and the Developing Countries* (pp. 292-315). Cambridge, England: Cambridge University Press.

Burfisher, M. E., Robinson, S., & Thierfelder, K. (2001). The Impact of NAFTA on the United States. *The Journal of Economic Perspectives , 15* (1), 125-144.

Burniaux, J.-M., & van der Mensbrugghe, D. (1991). Trade policies in a global context: Technical specifications of the rural/urban-north-south (RUNS) applied general equilibrium model. *OECD Development Centre Working Paper No. 48* .

Burniaux, J.-M., Delorme, F., Lienert, I., & Martin, J. (1988). Quantifying the Economy-Wide Effects of Agricultural Policies: A General Equilibrium Approach. *OECD Economics Department Working Paper No. 55* .

Canova, F., & Ortega, E. (2000). Testing Calibrated General Equilibrium Models. In R. Mariano, T. Schuermann, & M. J. Weeks (Eds.), *Simulation-Based Inference in Econometrics: Methods and Applications* (pp. 400-436). Cambridge University Press.

Costinot, A., & Rodríguez-Clare, A. (2014). Trade Theory with Numbers: Quantifying the Consequences of Globalization. In G. Gopinath, E. Helpman, & K. Rogoff (Eds.), *Handbook of International Economics* (Vol. 4). Elsevier.

De La Cruz, J., & Riker, D. (2014). The Impact of NAFTA on U.S. Labor Markets. *USITC Working Paper No. 2014-06-A* .

Deardorff, A. V., & Stern, R. M. (1986). *The Michigan Model of World Production and Trade.* Cambridge, MA: The MIT Press.

Deardorff, A., & Stern, R. M. (1981). A disaggregated model of world production and trade: An estimate of the impact of the Tokyo Round. *Journal of Policy Modeling , 3* (2).

Deardorff, A., & Stern, R. M. (1979). An Economic Analysis of the Effects of the Tokyo Round of Multilateral trade Negotiations on the United States and the Other Major Industrialized Countries. *U.S. Government Printing Office, Washington. MTN Studies , 5*.

Dixit, A. K., & Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *American Economic Review , 67* (3), 297-308.

Dixon, P. B., & Rimmer, M. T. (2002). *Dynamic General Equilibrium Modelling for Forecasting and Policy.* North Holland Pub. Co.

Dixon, P. B., Parmenter, B. R., Ryland, G. J., & Sutton, J. (1977). ORANI, A General Equilibrium Model of the Australian Economy: Current Specification and Illustrations of use for Policy Analysis. *2*.

Eaton, J., & Kortum, S. (2002). Technology, Geography, and Trade. *Econometrica , 70* (5), 1741-1779.

Fox, A. K. (2000). *Evaluating the Success of a CGE Model of the Canada-U.S. Free Trade Agreement.* Ph.D. dissertation, University of Michigan.

Francois, J. F., & Shiells, C. R. (Eds.). (1994). *Modeling Trade Policy: Applied General Equilibrium Assessments of North American Free Trade.* Cambridge University Press.

Ginsburgh, V., & Keyzer, M. (2002). *The Structure of Applied General Equilibrium Models.* Cambridge, MA: The MIT Press.

Gould, D. M. (1998, First Quarter). Has NAFTA Changed North American Trade? *Economic Review of the Federal Reserve Bank of Dallas* .

Hanoch, G. (1978). Polar Functions with Constant Two Factors-One Price Elasticities. In M. Fuss, & D. McFadden (Eds.), *Production Economics: A Dual Approach to Theory and Applications.* Amsterdam: North-Holland.

Hanoch, G. (1975). Production and Demand Models with Direct or Indirect Implicit Additivity. *Econometrica , 43*, 395-419.

Harris, R. G., & Cox, D. (1984). *Trade, Industrial Policy, and Canadian Manufacturing.* Toronto: Ontario Economic Council.

Head, K., & Mayer, T. (2014). Gravity Equations: Workhorse,Toolkit, and Cookbook. In G. Gopinath, E. Helpman, & K. Rogoff (Eds.), *Handbook of International Economics* (Vol. 4, pp. 131-195). Elsevier.

Hertel, T. W. (1997). *Global Trade Analysis.* Cambridge University Press.

Hertel, T. W., Peterson, E. B., Preckel, P. V., Surry, Y., & Tsigas, M. E. (1991). Implicit Additivity as a Strategy for Restricting the Parameter Space in CGE Models. *Economic and Financial Computing , 1* (1), 265-289.

Johansen, L. (1960). *A multi-sectoral study of economic growth.* Amsterdam: North-Holland.

Jomini, P., Zeitsch., J. F., McDougall., R., Welsh., A., Hambley., J., Brown., S., et al. (1991). *SALTER: A General Equilibrium Model of the World Economy: Model Structure, Database and Parameters.* Industry Commission, Canberra.

Kehoe, T. J. (2005). An Evaluation of the Performance of Applied General Equilibrium Models of the Impact of NAFTA. In T. J. Kehoe, T. N. Srinivasan, & J. Whalley (Eds.), *Frontiers in Applied General Equilibrium Modeling.* Cambridge University Press.

Krueger, A. O. (1999, December). Trade Creation and Trade Diversion under NAFTA. *NBER Working Paper No. 7429* .

Krugman, P. R. (1979). Increasing returns, monopolist competition, and international trade. *Journal of International Economics , 9*, 469-479.

Krugman, P. R. (1980). Scale Economies, Product Differentiation, and the Pattern of Trade. *American Economic Review , 70*, 950-959.

Larch, M., & Yotov, Y. V. (2016). General equilibrium trade policy analysis with structural gravity. *WTO Working Paper ERSD-2016-08* .

Leontief, W. (1966). *Input-Output Economics.* New York, NY: Oxford University Press.

Markusen, J. R. (2002). *Multinational Firms and the Theory of International Trade.* MIT Press.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica , 71* (6), 1695-1725.

Nicita, A., & Olarreaga, M. (2006). Trade, Production and Protection 1976-2004. *World Bank Economic Review , 21* (1).

Powell, A. A., & Snape, R. H. (1993). The Contribution of Applied General Equilibrium Analysis to Policy Reform in Australia. *Journal of Policy Modeling , 15* (4), 393-414.

Pöyhönen, P. (1963). A Tentative Model for the Volume of Trade between countries. *Weltwirtschaftliches Archive* (90).

Romalis, J. (2007). NAFTA's and CUSFTA's Impact on International Trade. *The Review of Economics and Statistics , 89* (3), 416-435.

Scarf, H. E. (1967). On the Computation of Equilibrium Prices. In W. Feller (Ed.), *Ten Economic Studies in the Tradition of Irving Fisher* (pp. 207–230). New York: Wiley.

Shikher, S. (2011). Capital, Technology, and Specialization in the Neoclassical Model. *Journal of International Economics , 83* (2), 229-242.

Shikher, S. (2013). Determinants of trade and specialization in the OECD countries. *Economic Inquiry , 51* (1), 138-158.

Shikher, S. (2012). Predicting the Effects of NAFTA: Now We Can Do It Better! *Journal of International and Global Economic Studies , 5* (2), 32-59.

Shoven, J. B., & Whalley, J. (1984). Applied General-Equilibrium Models of Taxation and International Trade: An Introduction and Survey. *Journal of Economic Literature , 22*, 1007-1051.

Shoven, J. B., & Whalley, J. (1992). *Applying General Equilibrium.* Cambridge University Press.

Shoven, J. B., & Whalley, J. (1974). On the computation of competitive equilibrium in international markets with tariffs. *Journal of International Economics , 4*, 341-354.

Tinbergen, J. (1992). *Shaping the World Economy: Suggestions for an International Economic Policy.* Applying General Equilibrium.

USITC. (1992). *Economy-Wide Modeling of the Economic Implications of a FTA with Mexico and a NAFTA with Canada and Mexico.* Washington, DC: USITC Publication No. 2508.

8 Appendix

Table 6 Policy-related trade barriers removed by NAFTA

Panel A Tariffs

Country	Food	Textile	Wood	Paper	Chemicals	Nonmetals	Metals	Machinery	Manuf.
Canada	8.83	17.65	8.48	3.46	8.26	7.78	4.83	5.63	8.51
Mexico	15.93	17.48	15.02	5.84	12.35	15.26	9.86	13.74	13.71
United States	2.14	10.64	2.47	0.62	4.48	7.43	3.04	3.37	4.68

Panel B Tariff equivalents of non-tariff barriers

Country	Food	Textile	Wood	Paper	Chemicals	Nonmetals	Metals	Machinery	Manuf.
Canada	3.23	7.95	12.96	0.00	1.23	0.00	11.82	0.87	3.33
Mexico	26.68	22.89	8.39	11.12	17.09	18.11	4.03	19.21	17.70
United States	11.07	5.81	2.63	0.67	3.28	0.51	0.00	4.05	4.10

Panel C Total policy-related trade protection

Country	Food	Textile	Wood	Paper	Chemicals	Nonmetals	Metals	Machinery	Manuf.
Canada	12.06	25.60	21.44	3.46	9.49	7.78	16.65	6.50	11.84
Mexico	42.61	40.37	23.41	16.96	29.44	33.37	13.89	32.95	31.42
United States	13.21	16.45	5.10	1.29	7.76	7.94	3.04	7.42	8.78

Table 7 Percent changes in import shares predicted by various models

Panel A GE gravity EK model

Importer	Exporter	Food	Textile	Wood	Paper	Chemicals	Nonmetals	Metals	Machinery
Canada	Mexico	36.3	188.2	21.0	14.9	23.7	21.1	16.0	65.6
Canada	U.S.	16.0	64.5	9.9	2.4	6.9	10.2	17.1	6.5
Mexico	Canada	-9.6	-36.3	-41.8	-24.9	-23.6	-22.9	-3.9	-2.8
Mexico	U.S.	19.6	21.1	2.1	4.0	13.7	25.5	9.0	18.7
U.S.	Canada	32.6	121.0	5.5	-2.8	30.5	26.5	11.3	49.0
U.S.	Mexico	107.1	337.9	60.1	37.2	73.8	69.0	28.9	129.1

Note: Each observation is a share of country *i* in country *n*'s imports of industry *j*.

Panel B GTAP model with standard GTAP elasticities

Importer	Exporter	Food	Textile	Wood	Paper	Chemicals	Nonmetals	Metals	Machinery
Canada	Mexico	28.2	147.9	13.6	9.1	19.8	12.4	36.4	47.6
Canada	U.S.	20.7	125.1	14.3	1.3	8.5	10.9	20.2	8.4
Mexico	Canada	24.5	32.6	2.5	0.7	13.3	21.9	10.8	24.1
Mexico	U.S.	33.2	37.3	10.0	6.2	18.0	32.7	17.6	28.8
U.S.	Canada	41.1	141.9	7.0	-1.6	17.2	23.0	3.7	28.5
U.S.	Mexico	60.6	176.0	14.0	11.7	34.7	35.8	24.7	80.8

Panel C GTAP model with all elasticities set to 8

Importer	Exporter	Food	Textile	Wood	Paper	Chemicals	Nonmetals	Metals	Machinery
Canada	Mexico	90.7	237.9	38.5	48.4	82.7	41.5	76.2	84.0
Canada	U.S.	33.7	142.4	17.8	3.2	16.8	15.8	26.6	10.1
Mexico	Canada	32.0	43.2	8.4	2.4	22.4	29.5	20.8	32.4
Mexico	U.S.	40.7	38.9	12.1	11.0	27.6	38.8	23.5	30.6
U.S.	Canada	77.1	190.3	13.3	-2.6	42.7	38.4	11.2	39.1
U.S.	Mexico	168.0	292.2	37.6	51.1	131.4	80.8	57.7	127.6

Panel D GTAP model with all elasticities set to 3

Importer	Exporter	Food	Textile	Wood	Paper	Chemicals	Nonmetals	Metals	Machinery
Canada	Mexico	28.0	72.3	19.4	15.4	24.4	17.7	27.2	23.6
Canada	U.S.	13.9	53.1	9.2	1.3	7.2	6.5	12.3	4.3
Mexico	Canada	23.3	24.0	5.7	3.6	14.4	20.1	10.1	18.4
Mexico	U.S.	25.0	23.7	6.6	5.5	15.4	21.8	10.5	17.6
U.S.	Canada	28.6	51.5	5.3	-0.3	16.6	15.2	4.7	15.5
U.S.	Mexico	46.4	70.0	16.0	15.4	36.3	29.0	18.9	35.7

Panel E GTAP model with GTAP elasticities, using Johansen's solution method

Importer	Exporter	Food	Textile	Wood	Paper	Chemicals	Nonmetals	Metals	Machinery
Canada	Mexico	20.2	85.1	12.3	6.9	14.8	10.1	26.1	30.2
Canada	U.S.	17.5	81.2	13.3	1.3	7.8	9.9	18.2	8.4
Mexico	Canada	25.7	34.7	4.8	1.7	14.3	23.8	11.3	32.7
Mexico	U.S.	30.1	38.9	9.9	6.2	17.5	29.3	16.1	35.6
U.S.	Canada	32.6	79.4	6.6	-1.3	15.5	20.4	3.9	26.2
U.S.	Mexico	41.6	90.0	11.8	9.5	26.6	28.2	19.2	52.0

Panel F Michigan (BDS) model

Importer	Exporter	Food	Textile	Wood	Paper	Chemicals	Nonmetals	Metals	Machinery
Canada	Mexico	10.0	11.0	22.7	15.0	-7.5	5.1	11.2	15.7
Canada	U.S.	7.4	24.3	3.4	3.0	2.2	3.8	2.9	1.6
Mexico	Canada	-7.0	5.6	12.1	0.8	7.0	156.0	6.2	14.4
Mexico	U.S.	7.4	10.5	11.3	2.0	2.2	7.1	5.7	10.8
U.S.	Canada	8.1	23.3	1.4	-0.1	2.8	58.9	8.0	4.2
U.S.	Mexico	10.3	14.7	4.4	4.4	-6.1	-23.6	14.8	41.9