A Moderate Compromise: Economic Policy Choice in an Era of Globalization
(Excerpt) - by Steve Suranovic (Palgrave McMillan, 2010)

Chapter 3: Why Empirical Data Can't Tell us What to Do about Policy

Economic theory strongly supports the following two conclusions: first, trade liberalization will result in a complex redistribution of income; and second, depending on the circumstances, trade liberalization can either lead to an improvement or a reduction in a nation's total welfare. Although theory can identify some of the characteristics of those who might gain and lose, without careful empirical measurement, it is impossible to know precisely **who** will gain and lose **how much** from freer trade. Whether national welfare rises or falls is important too, since economists often assuage the income redistribution concerns by suggesting that compensation be given to the losers from trade, drawn from the extra benefits accruing to the winners. However, the only way to assure that everyone gains from free trade after compensation is if there is a positive national welfare effect from trade liberalization.

Economic theory does not provide incontrovertible support for the proposition that free trade is national welfare enhancing, however. Although many theoretical models highlight the improvements in economic efficiency likely to arise when countries trade more freely, many other models provide arguments for why free trade may not be the best policy when there are market imperfections and distortions present. Because every country has a highly complex and unique mix of these imperfections, and because these will differ from country to country, theory can only reveal that the set of optimal policies will also differ from country to country. Only by analyzing the empirical data in particular countries would it be possible to determine what that set of policies might be,

or to determine if trade liberalization will be beneficial for any particular country. All of which points to the need for empirical measurement.

Simple Empiricism

One way to apply data to the theory is with what might be called *simple empiricism,* which uses empirical observations to suggest tendencies or inclinations. For example, suppose the trade data indicates that a country currently exports agricultural goods and imports electronic goods. Suppose collected data also shows that this country has many more unskilled workers per unit of capital than most of its trading partners. Using this broad aggregate data we can match the factor proportions theory to the data and suggest that trade liberalization is likely to reduce real income for individuals in the import-competing electronics sector but will raise real income for individuals in the exporting agricultural sector, at least in the short run. In the longer run we could say that the country's relatively scarce unskilled workers would experience income improvements while the country's relatively scarce owners of capital may experience income reductions.

For analytical simplicity, economic theory usually considers the effects of one policy change (such as trade liberalization) while assuming all other variables in the economy remain fixed at their original values. This means that while trade liberalization may cause import competing industries to lose income, this is only assured in the model when no other income improving economic effects are also occurring. In a real world situation many other changes are always taking place simultaneously though. Changes such as technological improvements, changes in management, outsourcing, and advertising effectiveness can all positively influence the outcome in a particular import

competing industry. If some of these changes occur simultaneously with the movement to freer trade, incomes in import competing industries might actually rise instead of falling.

For these reasons, theory and simple empiricism only suggest tendencies rather than predicting outcomes. We can reasonably say that trade liberalization will *tend to cause* - or that the probability is higher for - incomes to fall in import-competing industries. However, this is different, and less definitive than saying, trade liberalization will *surely* cause incomes to fall in import-competing industries. Unfortunately, theory combined with simple empiricism is never sufficient to predict actual outcomes.

Sophisticated Empiricism

The alternative approach is what we might call "sophisticated empiricism," by which I mean the use of econometric and statistical methods to search for correlations or cause and effect relationships or to develop empirical simulation models of the economy.

Using multiple regression analysis and other advanced techniques, it is possible to analyze whether, and to what extent, a whole collection of variables may have on a particular objective. For example, trade liberalization may be only one among many changes that could have influenced wages, or growth rates, or productivity. Other pertinent variables might be labor force and capital stock growth, interest rates and R&D spending, among other things. Using regression techniques it is possible to assess the influence of trade liberalization while simultaneously taking into account variations in other variables that may have also affected the same outcomes. If a positive effect is found, one can more convincingly argue that trade liberalization has indeed raised wages, growth rates, or productivity levels.

Another common empirical research approach involves the construction of what are called computable general equilibrium models (CGEs), also known as applied general equilibrium models (AGEs). These approaches begin with standard trade models, such as Heckscher-Ohlin-Samuelson (HOS), and introduce numerical estimates of their parameter values. Using a computer simulation it is possible to introduce a policy change, such as trade liberalization, and use the model to predict estimates for production, consumption, trade changes, employment changes and wage and welfare impacts on various groups in different industries. The models can show, at least at a high level of aggregation, which industries will benefit from free trade, which will lose and whether the net national effects will be positive or negative.

These empirical studies complement economic theory by quantifying the cause and effect relationships and by providing some evidence regarding the validity of the theories. Although researchers know that empirical evidence cannot *prove* the validity of theories, there is certainly widespread belief that if a theory or relationship cannot be supported by at least one or several instances of empirical evidence attesting to it, then the theory becomes suspect. In other words consistency of the data with the theory is a necessary, though not sufficient, condition to be believed.

Interestingly, in practice, failure of a theory to match empirical evidence does not necessarily lead to the refutation of that theory. An example of this was the discovery by Wassily Leontief in the 1950s that the US, a capital-abundant country, did not export capital-intensive goods as predicted by the HOS model. The "Leontief paradox," as it became known, did not lead researchers to throw away the model. Instead they developed ways of explaining why either the model was incomplete, or the data was

inadequate. Thus, theory refutation inspired creative activities to reconcile the data with the theory. In many instances this process generates a much richer understanding of the variables that influence outcomes and the problems associated with measuring those variables.

The Empirical Answer

Despite some well-known problems regarding the inability of models to convincingly verify theories, researchers and others do use empirical studies to support policy positions. Indeed, most of the arguments made by free trade advocates and opponents rely on numerous empirical studies that support the policy prescriptions being advanced. The list of studies lending direct or indirect support to trade liberalization is extensive. Here are just a few recent examples.

One recent example of direct support is a survey of a range of studies by Bradford, Grieco and Hufbauer (2006) that suggests that trade opening since World War II has added at least $800 billion to the US economy. They conclude that future trade liberalization can raise US incomes and standards of living. Other studies look indirectly at specific results that seem likely to correspond to generally good outcomes. Thus, Trefler (2006) provides empirical support that tariff reductions between the US and Canada resulted in sizeable increases in labor productivity. Since higher labor productivity should tend to raise wages and incomes and inspire GDP growth, the study supports the view that trade liberalization is good. Edmonds and Pavcnik (2005) provide a third example by suggesting that trade liberalization can result in less child labor usage in developing countries. Although they do not say that it *will* lead to less child labor,

their evidence that a higher incidence of child labor is not an automatic outcome of trade liberalization softens concern on that point.

Nevertheless, many other empirical studies lend support to free trade opponents. For example Traca (2004) shows that trade liberalization causes a decline in real wages and welfare of unskilled workers. Parikh (2006) suggests that liberalization may lead to an unsustainable balance of payments position. Finally Grieben (2005) suggests that southern-originated trade liberalization can result in an increase in Northern wage inequality.

Many of the books about globalization today involve lengthy arguments about why the "evidence" supports the author's point of view and why any counterarguments are notably weaker. The purpose of this chapter is not to undertake a comprehensive evaluation of the empirical literature or its ability to verify theories. Instead, I will make several simple observations about the reliability of empirical studies with regards to the policy debate of trade liberalization.

The relevant multi-part question with respect to trade liberalization policies is: can empirical studies, with a reasonably high degree of reliability, either individually or in combination, identify who benefits and who loses from trade liberalization over time AND show whether the net national effects are positive? A more general policy question is, can the optimal set of economic policies for an economy be identified with a high degree of reliability using empirical methods? Finally, rather than requiring a strong result that the policy be optimal, we can ask a weaker question, namely, can we use empirical results to determine which policies will make things better for a country?

Computable General Equilibrium Models (CGEs)

In the last chapter we discussed how to identify the optimal economic policies in the presence of multiple imperfections and distortions. This question is an extremely difficult one to answer because of the complex market interactions both within a country and internationally. Because of market interconnectedness and the widespread distribution of market imperfections, theory suggests that every policy that is implemented or changed by a government will generate a complex ripple effect of both positive and negative changes across the economy both immediately and over time. The overall impact of any policy change can only be identified if we can effectively measure these effects now and in the future.

An even more difficult question to answer is what the set of policies would be to optimize the well being of people within an economy or around the world. In other words, what are the set of policies that will maximize a country's well being? Economic theory shows that the answer to this question needs to account for the fact that every policy change affects the optimal policy level of every other policy within the economy. For example, if you reduce tariffs, then the optimal environmental policy, the optimal labor policy, the optimal competition policy, and all other policies will change as well. In principle then, one cannot evaluate the best trade policy in isolation from all of the other domestic policies that are already in place.

Probably the best and most direct way to account for these cross-market interactions is to construct a model that mimics these complexities as closely as possible. The closest economists have come to developing such a model is the computable general equilibrium (CGE) model. A general equilibrium (GE) model incorporates the

simultaneous interaction between goods markets and factor markets across at least several industries. GE models capture the connections of households, firms, governments and the foreign sector operating through different markets. Thus, in a general equilibrium model, a change affecting one market, will have ripple effects throughout the rest of the economy. A *computable* general equilibrium model quantifies the interindustry relationships using actual data collected from real economies.

The starting point for a CGE model is an input-output (I-O) table for an economy. An I-O table separates production into multiple sectors or industries such as agriculture, forestry, minerals, manufacturing, services, etc. The number of sectors specified can range from a handful to over a hundred depending on what data are available. An I-O table presents measured values for the amount of production from each industry that is used as an intermediate input into every other industry. In addition the table provides final good production from each sector. An I-O table is appended with information on labor and capital inputs in industries, demand patterns between consumption, investment, government and the foreign sector, among other info to create a social accounting matrix, or SAM.

Variations of CGE models may expand the number of labor inputs using data about the number of workers with different skills employed in different industries. On the consumer side, a CGE model can specify households with different income levels and consider variations in consumer preferences and income distribution. Typically the models assume perfect competition but adjustments can be made to allow for economies of scale in certain industries. Numerous such adjustments and extensions have been made as CGE models have been further developed and improved over the years.

The values in a national input-output table for a particular economy are based on the measured values in one year. Because it is difficult to compile all of these numbers, I-O tables are typically constructed only once every five years. Also, since the purpose of a CGE model is to consider how an economy will adjust after policy changes are made, additional information about the supply and demand elasticities is needed. Elasticities measure how demand and supplies change in response to variations in prices. Estimates of these elasticities are typically drawn from other studies in the literature.

Once constructed and quantified, the models are used to simulate the effects of changes in some of the key parameters. For example, tariffs on imported goods can be lowered, or eliminated, to simulate the effects of free trade area formation. With a detailed CGE model, the ripple effects on production, consumption, and government spending levels throughout the economy are numerically specified. Thus the models will indicate the magnitude of the effects in different segments of the economy. They can also be used to assess the overall national welfare effects of policy changes.

CGE models have become very popular in assessing the effects of trade liberalization for several reasons. First, these models are the best available for predicting changes in industry production levels and the number of jobs in a sector. Since much of the political discussion about trade liberalization often focuses on job effects, these models contribute to that debate. Second, CGE models provide the most complete evaluation of effects arising from economic policy changes. Thus, it is easy to sell these models to users in policy circles on the basis of sophistication. On the other hand, this sophistication can act as a drawback since it is sometimes difficult to explain their precision. There is also skepticism that because of the complexity of the models,

researchers could plausibly vary parameter values to produce virtually any result a policymaker might desire.[i]

Nevertheless, despite the complexity of these models and the hard work and expertise that goes into creating them, it is reasonable to ask how close they come to answering the questions we'd like to answer. In other words, can CGE models be used to convince us that trade liberalization is a good thing in some general aggregate sense? Can CGE models be used to identify the winners and losers from trade liberalization so that an appropriate compensation can be made? Unfortunately the answer to these questions is no; not now, not in the near future, and probably not even in the distant future will these models provide convincing answers to these critical questions. What follows are a few reasons why.

Aggregation

First, CGE models can identify winners and losers, but only at a high level of aggregation. CGEs are currently unable to identify income changes at finer levels of detail. The industry classifications consider industries like "manufacturing." For the manufacturing industry the data reveals how much intermediate and primary inputs come from various sources. For example, the data might show that for $1 of manufactured goods output, $0.02 came from the agricultural industry, while $0.31 came from the service industry, etc. The data would also indicate the value of capital and how many workers were employed in the industry, which in turn is used to identify productivity values.

In reality though, the manufacturing industry consists of production processes as diverse as semiconductors, clothing, food and beverages, machine tools, automobiles and

airplanes. At the micro level, the intermediate input proportions and capital and labor productivity will also vary greatly within the industry. Thus, when trade liberalization occurs, it can be expected to affect different manufacturing industries differently. Perhaps one manufacturing business using more capital will expand, while another using more services will contract. This would in turn change the input proportions and productivity values that, in the empirical model, are assumed to be fixed at the level of all manufacturing. When policy changes are made, a CGE model assumes that these aggregate industry relationships are maintained. In other words, manufacturing output will always require the same proportion of agricultural and service industry inputs and will maintain the same productivity of capital and labor.

Despite these problems, aggregation is necessary for two reasons. First, despite the advances in computing capabilities, CGE models take considerable time to solve. Computing requirements rise exponentially as production and consumption agents are separated into finer details; thus, the more aggregated the data, the less computation time and the easier it is to compute a solution. Second, aggregation is necessary because disaggregated data is simply not available across an entire economy. Although more data are beginning to be collected for some industries, it is expensive to collect detailed data describing the production and consumption processes. At best surveys can be done for some businesses and households and inferences drawn about the rest of the population. But this requires strong presumptions about regularities across a diverse economy.

Parameter Estimates

A second shortcoming of CGEs is that they assume that economic adjustments in the future will mimic patterns displayed in the past. This assumption is implicit because

all parameter values are based on past data. The I-O coefficients are usually updated only once every 5 years, if that. For some countries, I-O tables may not exist or there may have been only one such table constructed in recent years. In some cases, creative methods, such as cross-entropy techniques, have been developed to use more recent economic data to update the social accounting matrices to better reflect the values that prevail nearer the time of interest.[ii] Of course, past data is the only data available. Nevertheless to carry this data forward to the future in simulations means that a very important assumption is being made; namely, that the parameter values are unchanging.

Unfortunately though, parameter values are quite likely to change over time because trade liberalization will change the composition of production within each aggregate industry. Thus within each aggregated sector some businesses will expand while other will contract in response to freer trade. These changes should affect the input proportions from other industries measured in the I-O tables. In addition, some parameter values may be drawn from studies that are five, ten or even fifteen years old. And yet CGE models presume the future will mostly conform to the patterns of the past.

Incomplete Coverage

Perhaps the biggest problem with CGEs is the failure to include many of the market imperfections and distortions that surely are present in real economies. CGEs generally assume that markets are perfectly competitive, that full employment of all resources, which includes workers, always obtains, and that there are few market imperfections present. The models will incorporate some of the policy-imposed distortions by simulating government tax and tariff collections and expenditures. Also, sometimes the model will focus on a particular issue, such as environmental problems or

economies of scale, and incorporate at least that one imperfection into the model. However, despite modest attempts to add realistic feature into the models, there will always remain numerous other imperfections and distortions that cannot be included due to the sheer complexity.

Ken Arrow recognized the exclusion of one important imperfection when he wrote,

> "The most important novel development in microeconomics in recent years has been the recognition (by economic agents and by economists) of asymmetric information as a basic element in economic interaction. … I think it fair to say that none of these developments have been reflected in CGE models. The reason is clear. Economists have not developed any successful way of going from the individual decisions and outcomes of small-group interactions to the economy as an interacting whole."[iii]

Information asymmetries are an important feature of the world that affects economic decision making at the micro level. Since it affects all micro-level agents, it will also affect macro-economic outcomes like those being predicted in CGE models.

But the problem is much more serious. It's not just information asymmetries that are missing. Most models are missing a complete and accurate representation of labor and capital market rigidities, environmental effects, strategic behavior in oligopolistic markets, social and cultural effects, positive and negative externalities, and public goods, among many others.

The theory of the second best teaches that the only proper way to assess policy impacts is to recognize all of the market imperfections and distortions simultaneously. To exclude one or more, means that policy evaluations will certainly be incomplete and thus more likely to be incorrect. Thus, to assess whether free trade is good or bad from a

national perspective, while using an incomplete model of the economy, should by no means be convincing to any policymaker or any objective observer.

Jagdish Bhagwati, a prominent economist who has contributed substantially to the trade literature, summed up a common professional impression of CGE models when he wrote, "in fact I consider many of the estimates of trade expansion and of gains from trade – produced at great expense by number-crunching at institutions such as the World Bank with the aid of huge computable models, and then fed into the public policy domain with the aid of earnest journalists – as little more than flights of fancy in contrived flying machines." (Bhagwati 2004, p. 230) Unfortunately, for policy prescriptions, a review of the problems inherent in computable CGEs suggests he is exactly right.

The Value of CGEs

If CGEs are viewed with such apprehension, we might ask whether they are of any usefulness whatsoever. The answer is a restrained yes.

Many times when policy discussions occur, participants in the debate often talk as if the policies will have a narrow effect only in the industry of interest. For example, if the steel industry argues for a protective tariff, they will emphasize the protection of jobs for workers, often as if that is the only relevant effect. A careful economic analysis of the issue, however, will point out the effects on consumers of steel products, the effects on other industries who use steel as an input, the effect on government revenues, the effects on income to different production factors, the effects on the environment and much more. Indeed, one of the important lessons of economic theory is recognition of the interconnectedness of markets. These relationships are nicely described using general equilibrium models.

CGEs put numbers into the theory to demonstrate the kinds of pattern changes that may occur from various policy changes. As such, they are helpful in highlighting the complexity of the potential effects. In other words, the simulations can depict how policy changes will ripple through the economy even affecting industries that appear quite remote from the steel sector.

Sometimes the results of CGEs show outcomes that the researcher can't easily explain because the process is hidden deeply in the complex mathematics. For some this becomes a source of criticism for these models, however I think these unexplainable outcomes are one of the main reasons CGEs are informative. Even though these models are nowhere near the complexity needed to mimic the real world, they still complicated enough as to be incomprehensible sometimes. And, if models much simpler than the real world can display policies with incomprehensible impacts, then what does that imply about our ability to understand the effects of *real* policy actions in *real* world economies? Surely it should strongly diminish any hope that we can confidently predict the specific effects of policy actions in the real world.

The results of CGE analysis will always display complex interactions, the general patterns of which will surely occur. However, the predictions of any CGE model will also surely be "flights of fancy." They should not be used to inform policy choices because they cannot answer the questions that need answering to be useful, namely, 1) who will gain income and who will lose from the policy change, and 2) will the nation be better off after the policy change? Without answers to these questions we cannot be sure that the policies will be good for the nation nor could we implement a compensation scheme to assure that everyone benefits.

Indeed, Kehoe (2005) conducted an evaluation of three of the most prominent CGE models from the early 1990s used to project the effects of the NAFTA. He found that all of the models drastically underestimated the effects on trade between the countries. The results of these models, though extremely sophisticated, should simply not be trusted.

Piecemeal Empirical Investigations

The popular perception in the economics field today is that economic theory requires empirical support to be believed. For this reason trade theories that suggest positive economic effects as a result of trade liberalization have been taken to the data, as it were, to see if the results can be verified. Numerous empirical studies exist in the literature that suggest positive outcomes from trade liberalization. Although all researchers acknowledge their limitations, they are also likely to believe that each study moves us closer to the truth. There is power in numbers. Most would agree that if only one study showed support for free trade then the proposition that free trade is good would be accepted only with great reservations. However, if two or three or a hundred studies using different data sets and different estimation techniques also show support for free trade, then the case for free trade strengthens substantially. Although studies opposing free trade may weaken the case for free trade, these can be countered with more analyses showing support. The impression is that continuing empirical research steps us closer and closer to the truth.

Based on the philosophy of science literature, especially that of Karl Popper and Imre Lakatos, most researchers accept that empirical data cannot be used to verify theories.[iv] Just because the data is consistent with a theory does not *prove* it would be so

in all instances. Instead it can only be said that in the instances tested, the data is *consistent* with the theory. An example of the potential problem is illustrated in Taleb (2007). Consider a simple theoretical statement or proposition: "all swans are white." If this proposition were tested empirically in Europe and the Northern hemisphere by looking extensively at bird species and noting the color of all swans, the theory would have been supported over and over again. However, once Europeans discovered Australia, the theory would fail upon the first sighting of the Australian black swan. One of Taleb's central propositions is that scientific knowledge is much less certain than it often appears and that empirical observation is not sufficient to prove very much.

The purpose of this section is not to review the literature in the philosophy of science, but rather to suggest a few reasons why the empirical literature tending to support trade liberalization makes a weaker case for trade liberalization than is commonly suggested. The main reasons highlighted are, a) partial analysis does not identify the total effects, b) many effects from trade liberalization are immeasurable, c) researcher confirmation bias can tarnish objectivity, and d) the Bayesian method is not applicable.

Some researchers may argue that empirical tests are not designed to inform policy debates, instead they simply represent tests of theories versus alternatives. In this view theories and tests of theories are designed to be positive economic analysis that simply try to explain "what is" rather than normative analysis attempting to support what "should be." Although this bifurcation makes some sense, it is unrealistic to think that public policy advocates are not using empirical results to build support for their preferred policies. Simply listen to any policy debate and take note of how many times "numbers" are used to support a position. Since even purely positive studies can be used in

normative discussions, it still makes sense to consider them in terms of their policy implications.

Partial Analysis

Because the question, will free trade promote an increase in the overall national welfare, is very broad and difficult to answer, researchers simplify the questions to make them more manageable. This is both reasonable and necessary. For example, a researcher may inquire whether trade liberalization in a group of countries has affected GDP growth or poverty, or wage inequality. Each of these is only a partial question however.

For example, while it is certainly true that GDP contributes to the well being of a nation, it is also well accepted that many things people care about are not captured in this measure. For example, GDP is a measure of the *production* of all goods and services during *one* year in a country. It does not account for income inequalities, does not include negative effects like pollution caused by some productive activities, does not capture benefits caused this year by previously produced products such as used cars and houses, and only imperfectly accounts for inflationary effects. Furthermore, GDP measures production not consumption. When a nation runs a trade deficit, its national consumption expenditures exceed its productive output and this contributes to a higher standard of living than represented by its GDP.

Nonetheless, because evaluating all of these impacts is very difficult, it makes sense to answer the simpler partial question. Thus, many studies have indicated that trade liberalization has indeed been associated with faster real GDP growth for many countries in the past. This result is used by advocates of free trade to suggest that trade

liberalization is a good thing. And it may well be a good thing. However, several studies using partial and incomplete indicators of well-being are insufficient to prove that trade liberalization raises national welfare. Skeptics can reasonably argue that GDP growth isn't everything people care about and just because GDP grows doesn't mean we should choose free trade. This counterclaim is just as plausible because we simply cannot know very much about overall effects from a partial analysis alone.

Immeasurability

One reason partial analysis of the effects of policy changes is necessary is because many of the effects from policy changes are simply immeasurable. Consider, as an example, the effects of trade and globalization on culture. Some argue that globalization leads to the decline of traditional industries. An early example of this effect was the decline of the Indian textile industry in the 1800s after the introduction of new weaving technologies and trade with Britain. The cheap textiles imported from Britain devastated the industry. In modern times one will notice the resistance that prevails in developed countries to liberalization of agriculture. In the US they talk of the destruction of the family farms. In Europe they fear the loss of the pastoral countryside and the rural reminders of a simpler time. In Japan they maintain agricultural traditions with rice paddies in the middle of urban areas.

Although one can argue that references to cultural traditions by supporters of agricultural restrictions are merely a rhetorical cover for their real concern, namely fear of losing income, it seems reasonable to imagine that cultural considerations can affect people's well being. Indeed these utility effects can be modeled as a negative externality, a kind of market imperfection, which can then justify the use of trade restrictions to

preserve culture. However, to implement such a policy appropriately would require the measurement of the effects of trade on culture and the subsequent influence on individuals utility or well being. As far as I know this has never been attempted, largely because the idea of culture itself is rather fuzzy and because enumeration of its effects, even if possible, would certainly be based on a considerable number of arguable assumptions.

If culture is accepted as a relevant factor in assessing the effects of globalization, and if it cannot be measured convincingly, then one can never assure people that freer trade is indeed a good thing. Negative effects on people from cultural changes could overwhelm other positive effects such as GDP growth. Indeed, perhaps the inability to measure cultural effects is a reason economists seem reluctant to accept it as a possible motivation for trade restrictions. Cowen (2002) is one economist who clearly accepts that culture matters. His book provides an extensive overview of the effects of globalization on culture around the world. Although he makes no attempt to measure the effects, he does argue that cultural changes are not all bad and that the positive effects of cultural change caused by globalization may overwhelm the negative effects. Of course we can never know if his argument is valid unless we could somehow measure culture convincingly.

A similar issue arises with the recent issue of global climate change. If energy usage rises because of increased economic activity inspired by globalization, then the increased carbon emissions into earth's atmosphere may be causing a dramatic shift in the earth's climate. Average temperatures are expected to rise across the planet, possibly causing changes in rainfall, increases in the number and strength of hurricanes and

typhoons, and the melting of large land ice masses, thereby causing a rise in the ocean level. Economic damage may be catastrophic, especially if a rising ocean buries coastal cities under meters of water, if superstorms become commonplace, and if changes in ocean temperatures influence regional climates. For example, there is some concern that rising ocean temperatures could stop the gulf stream that carries warm water towards Europe. If the gulf stream stopped, Europe would become much colder, greatly affecting agricultural production and many other economic activities.

Indeed significant global climate change would surely have important and perhaps catastrophic effects on the world's economy. But how much of an effect is an open question. In fact, the issue is so complex that the extent of the damage is impossible to assess. Perhaps this is one reason some people are unwilling to accept climate change as a real phenomenon, believing instead that it represents a left-wing conspiracy to wrest control of people's lives.[v]

Of course, difficulty with measurement won't prevent researchers from making assessments. Nevertheless, consumers of this information need to be aware that all assessments of such far reaching and global phenomenon will be based on an enormous number of assumptions, some of which will certainly turn out to be invalid. In the same way as CGE analysis, assessments of global climate change can be helpful in providing a "sense" of the kinds of outcomes that are in the range of possibilities. However, also like CGE studies, they are quite likely to be wrong in their details *ex post*.

Confirmation Bias

One well-known problem in empirical research is known as confirmation bias. This occurs when a researcher begins with an expectation that a particular result is true

and then searches selectively for evidence to support the expectation. Some evidence of this problem seems apparent in econometric testing.

For example, most trade economists generally believe that free trade is a good thing for a country. A search of the empirical literature on the effects of trade liberalization finds many more studies tending to support free trade and few tending to oppose it. One possibility to explain this pattern is that the actual evidence more strongly supports the positive effects of trade liberalization. In this case no confirmation bias exists. However, it may also be true that results tending to support trade liberalization are more likely to be published in trade publications since the reviewers of these papers will also be economists who tend to believe in the net benefits of free trade. If empirical researchers recognize that the bias exists, they may be inclined to only submit papers that have the expected results.

Observation of the standard operating practice in economics may support this claim. For example suppose a researcher decides to study the empirical effects of trade liberalization on poverty rates. Initially, a model will be used to identify all of the variables that may influence poverty in a country and the way in which these variables may interact. The model will be used to inform the specification for an econometric test and data will be collected for the test.

Quite frequently, when the test is run the first time, the results are mixed. The initial hypothesis, that trade liberalization reduces poverty, may be only weakly supported or perhaps even rejected. At this point the researcher has a choice. Either she can attempt to publish these "weak" results, or she could check to see if a stronger result might be obtainable.

Suppose however that this researcher "believes" that that trade liberalization should reduce poverty. In this case, she will be inclined to also believe that something was wrong either with the model specification or with the data. In all but the simplest empirical studies there are always a variety of "problems" with the data. For example, rarely is the available data for a study precisely the data a researcher would wish to have. Thus, proxy data is used as a substitute.[vi] The proxy data may or may not mimic the behavior of the ideal data that was unobtainable. Standard operating procedure in the profession is to rerun the study, sometimes hundreds of times, with different data and different model assumptions. Indeed, most of the development of econometrics involves improvements in techniques to compensate for the myriad of data problems that can arise. If alternative specifications with alternative data provide greater support for the hypothesis, then it is generally believed that this new specification and data is "correct." This process is sometimes called *data mining*, or *massaging the data* to secure a better fit with expectations.

Unfortunately though, it is impossible to know whether the final model specification and data proxies are closer or further from the truth. This would be especially true when the researcher has a preference about the final result. If the researcher approached the problem completely objectively, he might conduct his research study very differently. For example, first he would fully investigate *ex ante* precisely the "best" data available and the "best" possible test of the hypothesis. Then, he would run the model *once* and publish whatever results were obtained. In this way, confirmation bias might be reduced.

Problems with confirmation bias are not unknown to empirical researchers in economics. For example, Leamer (1983) provides suggestions on how to improve the empirical procedures with an appropriate use of sensitivity analysis that considers a range of studies that vary with different assumptions about the prior beliefs of the researcher. The example he offers considers the deterrent effects of the death penalty. He shows how a researcher using one set of assumptions can reach a conclusion that the death penalty reduces homicides. However, by altering the model, either including or excluding variables deemed important or not, he is able to show that some equally plausible models shows the death penalty actually raises the homicide rate.

Procedures involving alternative tests such as these have become common practice in economics. In order to have a paper published today in the best economics journals, most research must show that a particular result is supportable under a variety of model specifications and with alternative data assumptions. These procedures surely act to reduce confirmation bias, but it is unlikely to eliminate it entirely.

Bayesian Inference

Most researchers would agree that one or two empirical studies supporting the argument for freer trade would not be too convincing. Nonetheless, most may also accept that if numerous studies, conducted by many different researchers, using many different data sets, over many different periods of time, all show support for trade liberalization, then the case is much more solid. Even with substantial evidence such as this, one should never claim that the point has been "proven." However, one can claim that the probability that trade liberalization is overall good for a country is certainly higher.

Strong believers in free trade may even believe that the probability free trade is good for a nation is very close to one hundred percent.

The idea that the probability of a hypothesis being true rises as more supporting evidence is accumulated is know as Bayesian inference or Bayesian updating. For example, suppose, perhaps based on economic theory alone, someone believes that free trade is good for a country with probability 40%. Afterwards, every empirical result the person reads that supports the original hypothesis increases his belief in free trade a little bit. Similarly, every empirical result that casts doubt on the proposition reduces the probability. If studies supporting free trade are greater in number and also more believable (perhaps because of the care and rigor of the analysis), then his impression that free trade is a good thing will grow.

While Bayesian inference is a perfectly valid principle, there are several problems in applying this process to the issue of trade liberalization and globalization. First, remember that trade theory teaches that trade liberalization will always cause benefits to some individuals and losses to others. Thus, a finding that some overall good will result from trade liberalization is not surprising. Furthermore, the good outcomes are likely to arise in the form of improvements in economic efficiency, which will translate into higher incomes and national GDP. The negative effects are more likely to arise because of market imperfections, which may involve effects like pollution, global climate change, cultural changes and national security concerns. These latter effects are much less easily measured than the efficiency effects.

Thus, if researchers do empirical studies using the most readily *available* data, and if that data is more likely to display the positive effects of trade liberalization, then

they are more likely to produce empirical studies tending to favor trade liberalization. Furthermore, if economists also harbor some bias in support of freer trade, then the studies they decide to pursue are also more likely to show support for trade liberalization. While Bayesian updating may lead people to believe more strongly in free trade, that perception could be because of a bias in the type of studies conducted and the type of data available.

A second problem with the Bayesian approach is that no one really knows how to precisely determine the probability values. Although it makes sense that the greater the number of studies favorable to free trade, the greater the probability that free trade is a good thing, it makes a big difference to policy makers if the increase is from a 40% to a 90% probability or from a 40% to a 43% probability. In the former case, if most people accepted the probability values, then the hypothesis that free trade is good would be accepted by a near consensus. In other words, empirical testing would be sufficient to convince most people what is most probably true. However, if the probability rises only from 40% – 43%, then empirical tests are not likely to change popular opinion. Also, since there are so many impacts from freer trade and globalization that are almost impossible to measure effectively, it seems unlikely that we could be anywhere near 90% (or even 80% or 70%) certainty that free trade is an overall good thing.

Finally, even if empirical testing were sufficient to convince most people that free trade was good for a country overall, that still is not sufficient to guarantee that everyone will enjoy the benefits of freer trade. Only after an appropriate redistribution will the gains be adequately reallocated so that all people may benefit form the net welfare improvement. But, redistribution requires identification of winners and losers in highly

complex economies. Unfortunately, empirical analysis has not come anywhere close to accomplishing this objective.

Conclusion

Is free trade ultimately the best policy for countries to pursue? The theoretical literature says that trade liberalization will surely generate a complicated mix of winners and losers over time and that the sum total effects are uncertain. The only way to determine the answers, presumably, is to go out and measure the effects. The fundamental questions that need answers are:

1) which individuals will gain and lose from freer trade, how much will they gain and lose, and when will they gain and lose?
2) will the sum of the benefits over time to those who benefit from trade liberalization exceed the sum of the losses to others?

If both questions can be answered affirmatively then the choice of policy is clear. Free trade coupled with an appropriate compensation scheme can raise everyone's well being in the future. However if we can't answer these questions then the standard cost-benefit approach to policy choice is hopeless.

Unfortunately, the conclusion of this chapter is that empirical analysis does not provide an effective method to decide what kind of policy actions to take regarding trade and globalization. Advanced empirical methods cannot tell us whether free trade is good for a country overall and they cannot tell us precisely who will win and lose from policy actions. Empirical methods also cannot tell us precisely what set of policies would be optimal for a country from a national perspective. This means empirical methods cannot

tell us what to do about policy. Free trade might be best for a country, but we cannot know for sure. Some combination of domestic and trade policies might be best for a country, but we cannot know with sufficient certainty what that set of policies would look like.

These conclusions follow despite the fact that empirical methods have improved significantly over the past century. Data collection and computing capacity continue to grow. The care and rigor applied in the best research in the discipline is impressive. However, despite these advances, the questions that can be asked and the phenomenon that can be measured effectively remain much too simplistic in comparison to the complexities that prevail in the real world. Thus, while a good study will provide the very best answer possible to the question being asked, that question by necessity will be extremely narrow in its scope. Indeed, even the full range of empirical studies from the beginning of the statistical era, are extremely narrow in scope. There are simply too many relationships that cannot be specified and too many effects that cannot be measured. It also seems unlikely, given how distant our current analytical state is from the complexities of the world, to expect that empirical analysis will provide an answer to these questions anytime soon.

Another perspective on this conclusion is possible with a simple analogy. We might ask whether the empirical investigations into the question of free trade and globalization are more like,

a) showing that the earth and planets revolve around the sun; or

b) showing whether the one sports team is better than another.

In the first case, Galileo's observations (empirical tests) that Venus displayed different phases, precisely like the earth's moon, was sufficient to convince most scientists that the heliocentric theory offered by Copernicus and others was valid. I don't know of anyone today who seriously doubts that the earth revolves around the sun. Thus, empirical observation and measurement was sufficient to answer this question definitively.

However, the question about which sports team is better is more difficult to assess. The first issue is how one should define the term "better." If the teams meet and team A defeats team B, some might find that sufficient evidence that team A is better. But what if the teams meet several times and each team wins at least once? Or what if they never play each other? Sports enthusiasts typically can rattle off reams of statistics attesting to the superiority of their favored team. However, in most instances, for teams at similar competitive levels, different observers can reasonably reach different conclusions. It will be impossible to reach agreement by most observers on which facets of the teams are most important and what it really means to be "better." Thus, for this relatively simple question, empirical observations are not likely to lead most observers to a definitive conclusion.

The conclusion of this chapter is that the question - is trade liberalization good for countries? - is more like deciding which sports team is best than it is like deciding whether the earth revolves around the sun. The question itself is unanswerable, no matter how many numbers we collect or how much computing capacity we employ. If this is indeed true, then the debate over trade liberalization may be endless. Indeed if we look back over the past several hundred years, many of the arguments both in favor and

against free trade have not changed very much. While it is true that more sophisticated mathematical analysis is often used to make the points today, it is not clear that this information is helping to move people or countries towards a consensus on the issue.

Indeed mathematical analysis may be more of a problem than a solution. One of the major advances in economics in the 20th century was to apply the scientific method to economic issues. Economics began to mimic the physical sciences. Economic relationships were increasingly described using mathematical models with the necessary assumptions carefully specified. Results in theory required confirmation via empirical testing. In this way, tests of economic relationships tended to mimic tests of physical relationships.

A similar scientific approach is now applied to policy evaluation. To decide whether a policy is appropriate, the researcher collects data for countries with variations in the policy and using an empirical test, determines if the policy had positive or negative effects. However, an implicit assumption is being made here that may not be valid. Physical systems display regularities that may not arise in social systems. For example, water is known to boil at 100°C at sea level. Given physical regularities, we expect water will boil at the same temperature today, tomorrow and 10,000 years from now. It will boil at 100° in the US, in China and even on Mars (assuming the same atmospheric pressure is maintained). Social systems may not display this same regularity. Thus, if free trade is shown to improve GDP in a sample of 50 countries today, it may not follow that free trade will improve GDP in another sample of countries 100 years from now. Changes in the behavioral patterns of participants in the system may prevent a replication.

The implication is that empirical methods in economics suggest a belief that economic and social relationships follow immutable patterns. With that belief, one may begin to see the economy like a large machine, whose parts interact with each other in discoverable ways and with the regularity of a physical system. Once we understand how the economic machine works, and once all of its effects have been measured, it becomes possible to use policy levers to produce more favorable economic outcomes. Afterall, policy changes are like changing the settings on the machine. Adjust the settings appropriately and one can make the machine work more effectively.

However, this scientific procedure is not very different from the socialist/central planning ideas of the past. Hayek (1988) deemed these intentions the *Fatal Conceit*; namely the idea that we could fine tune a large macro economy using policy levers so as to produce an outcome that would be better in some overall sense. In a similar vein empirical tests used to support freer trade presume that by measurement we will determine that free trade is best and then implement the policy choice together with an appropriate redistribution to assure that all will benefit. Hayek argued that central planning could never work and instead we should succumb to the workings of the private market. In a similar vein I argue that current empirical methods used in policy evaluation are not effective or convincing, therefore it is imperative to find an alternative choice mechanism.

---

[i] Kehoe, Srinavasan, and Whalley (2005; p.9)

[ii] See Robinson, Cattaneo, and El-Said (2001).

[iii] See Arrow (2005).

[iv] See for example Lakatos and Musgrave (1970) and Popper (2002).

[v] For example, see the remarks by Vaclav Klaus, President of the Czech Republic, at the UN in 2007. http://www.globalwarmingheartland.org/article.cfm?artId=22021

[vi] Proxy data is an alternative data series that are presumed by the researcher to be highly correlated with the preferred data series. However, since the preferred data series is not available it is also impossible to prove that a strong correlation exists.