

University of Chicago Law School

Chicago Unbound

Journal Articles

Faculty Scholarship

1998

Comparative Advantage and the Normative Economics of International Trade Policy

Alan O. Sykes

Follow this and additional works at: https://chicagounbound.uchicago.edu/journal_articles



Part of the [Law Commons](#)

Recommended Citation

Alan O. Sykes, "Comparative Advantage and the Normative Economics of International Trade Policy," 1 *Journal of International Economic Law* 49 (1998).

This Article is brought to you for free and open access by the Faculty Scholarship at Chicago Unbound. It has been accepted for inclusion in Journal Articles by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.

COMPARATIVE ADVANTAGE AND THE NORMATIVE ECONOMICS OF INTERNATIONAL TRADE POLICY

*by Alan O. Sykes**

INTRODUCTION

The law of international trade, national and international, governs the manner in which governments intervene in the free flow of goods and services across borders. In this inaugural issue of the *Journal of International Economic Law*, it is especially appropriate to review the state of economic learning that bears on the wisdom of such intervention. Much of what I have to say will be familiar to economists, and my goal in this paper is not to break new ground but to convey the key insights from the normative theory of international trade in an accessible manner that presupposes no economic training beyond the capacity to appreciate a simple supply and demand diagram.

I will begin with the economic theory relating to the impetus for trade, explaining the economic concept of 'comparative advantage', suggesting why the existence of comparative advantage creates profitable opportunities for international exchange, and discussing the possible sources of comparative advantage. I then consider the normative economics of trade based on the exploitation of comparative advantage, reviewing the case for free trade and its caveats, as well as the economics of trade discrimination and of the choice among protectionist policy instruments. A future contribution will survey the positive theory of international trade policy and trade agreements.

The prose is cast largely with reference to trade in goods, but I do not wish to slight the increasingly important area of trade in services. Most of what I have to say applies to trade of any sort, and requires at most minor modification for services sectors.

1. THE ENGINE OF TRADE: COMPARATIVE ADVANTAGE

International trade occurs because a buyer in one country desires something produced in another country, and is willing to pay the price required to obtain it. Implicitly, the buyer in question must prefer the imported item to a domestically produced substitute, either because it is cheaper or of higher quality (or both), or because domestically produced substitutes are

* Frank & Bernice J. Greenberg Professor of Law, University of Chicago. The author is grateful to Alan Deardorff for thoughtful comments.

unavailable. The theory of comparative advantage affords the predominant explanation for why such circumstances arise.¹ We begin with the meaning and consequences of comparative advantage, and will then consider its genesis.

A simple illustration

Like many international economics texts, I will develop theory of comparative advantage in a simple, numerical illustration. Given the simplifying assumptions necessary to this illustration, the reader may wonder whether it has any generality or real-world applicability. In fact, none of the assumptions made here is logically necessary to anything of importance, and they merely serve to facilitate an accessible exposition. The next section indicates how greater generality on all fronts makes no essential difference.

Thus, consider an exceedingly simple global economy, with only two nations, A and B. Each nation has its own labour force, and let us assume that it is impossible (or unattractive) for labour to migrate from one nation to the other. The only input into the productive process is labour (measured in units of time), and all workers are identical. The only outputs are ‘guns’ and ‘butter’. It is perhaps useful to think of this economy as one without firms, where the workers in each nation must simply choose whether to allocate their labour to gun or butter production. All markets are competitive.² Let transportation costs for guns and butter between countries be zero. The unit of currency in country A is the \$, while in country B it is the £. Lastly, let production in each nation occur in accordance with the following input–output table:

Labour Requirement Per Unit of Output		
	Guns	Butter
Country A	1.0	2.0
Country B	2.0	3.0

From an examination of the input–output table, observe that gun production in country A requires only half as much labour per unit of output as in country B, while butter production in country A requires only 2/3 as much labour per unit of output as in country B. Accordingly, country A has *absolute advantage* in the production of both guns and butter – country A is

¹ The theory of comparative advantage is developed at length in any international economics textbook. Excellent basic texts include Peter Kenen, *The International Economy* (1985); Paul Krugman & Maurice Obstfeld, *International Economics: Theory and Policy* (1994); and Peter Lindert, *International Economics* (1991). More advanced texts include Avinash Dixit and Victor Norman, *Theory of International Trade* (1980), and Jagdish Bhagwati and T.N. Srinivasan, *Lectures on International Trade* (1983).

² By this I mean that in equilibrium, the market price of any good or service is equal to its marginal cost.

better at everything in this simple economy. One might thus be tempted to conclude that country A will have no interest in trading with country B. But this conclusion would be incorrect, for despite its absolute inferiority in all lines of production, country B nevertheless has *comparative advantage* in the production of butter, and can export it profitably.

To see why, we begin by asking what the prices will be for guns and butter, in each country, in the absence of international trade (so-called *autarky*). The assumption that markets are competitive implies that each good will sell, in each country, for its marginal cost of production. The marginal cost of each good is simply the cost of the number of units of labour that go into it. With no loss of generality, assume that the currency units in each country are such that the market price of a unit of labour in autarky is 1.0. Hence, the autarky prices for guns and butter in each country will be equal to their labour input requirement. We can thus modify the input-output table slightly to create a table of autarky prices:

	Guns	Butter
Country A	\$1.00	\$2.00
Country B	£2.00	£3.00

These prices accord with common sense: If it takes twice as much labour in country A to produce a unit of butter as it does to produce a gun, then a unit of butter ought to be twice as expensive. Further, if the market price of the labour to produce a gun is \$1.00 and if prices reflect their marginal costs, then a gun should cost \$1.00 and a unit of butter \$2.00. Equivalent reasoning produces the respective prices of £2 and £3 in country B.³

From this starting position of autarky, imagine that an entrepreneur from country A visits country B, and happens to bring along a gun. The entrepreneur observes the market prices for guns and butter in country B, and comes to the realization that the gun can be sold for enough local currency (£2) to buy $\frac{2}{3}$ of a unit of butter. The butter can then be transported back to country A (at zero cost given my earlier assumption) and sold at a price of \$1.33 ($= \frac{2}{3} \times \2.00). The returning entrepreneur can then buy a new gun in country A for \$1.00, and still have \$0.33 left over as profit for the transaction. He will quickly realize as well that by expanding the scale of operation, exporting lots of guns and importing lots of butter, a good deal of money can be made.

Had the entrepreneur from country A brought butter to country B rather

³ Note that nothing turns on the assumption that the price of labour in each country is 1.0 in the local currency; if the price of labour were greater or less than 1.0, the autarky prices for guns and butter would simply be scaled up or down by the same proportion and everything to be said below would remain true after an appropriate change of units.

than a gun, however, no such profit-making opportunity would exist. A unit of butter fetches only £3 in country B, which buys only 1.5 guns (double everything if the notion of 1/2 gun is bothersome). The 1.5 guns can be sold in country A for \$1.50, which is \$0.50 shy of what is needed to replace the unit of butter that was sold in country B to get the 1.5 guns – the transaction thus loses \$0.50.

The analysis works in reverse if we imagine that an entrepreneur from country B visits country A and brings along some butter. The reader can readily verify that selling butter in country A at the autarky price, buying guns with the currency earned on the sale, and then returning to country B to sell the guns, is a profitable venture. Likewise, it is not profitable to bring guns into country A for the purpose of selling them and converting the currency into butter for sale in country B.

Entrepreneurs from country A can make money selling guns in country B because country A has *comparative advantage* in the production of guns. The entrepreneur from country B can make money selling butter in country A because country B has *comparative advantage* in butter production. Perhaps the easiest way to understand the concept of comparative advantage is to restate the autarky prices for each good in terms of the foregone production of the other good that is necessary to produce one unit of the good in question (the ‘opportunity cost’ of a unit of production in terms of the other good). Thus, in country A, because a unit of butter production requires two units of labour that could have been used to produce two guns, the price of butter in terms of guns is 2.0. Reciprocally, the price of guns in terms of butter is 1/2, because a reduction of butter production by one-half unit frees the labour necessary to produce one gun. In country B, the analogous reasoning implies that the price of guns in terms of butter is 2/3, while the price of butter in terms of guns is 1.5. When these prices are compared, it is evident that country A has the lower price of guns in terms of butter – 1/2 versus 2/3. Country B has the lower price of butter in terms of guns – 1.5 versus 2.0. Accordingly, gun production in country A sacrifices fewer units of butter production than it does in country B, and butter production in country B sacrifices fewer units of gun production than it does in country A. One can thus say that gun production is *comparatively* more efficient in country A, and that butter production is *comparatively* more efficient in country B.

These comparative efficiencies, as has already been shown in this example, are all that is necessary to create the opportunity for profitable international trade. They beget a difference across nations in the *ratios* of the prices for goods sold in autarky, which entrepreneurs can exploit by exporting the good that is relatively cheap locally (in terms of the other) and importing the good that is relatively expensive locally (in terms of the other).

The theory of comparative advantage thus yields a simple prediction: nations will tend to specialize in the production of goods in which they have

comparative advantage, exporting them to other nations in exchange for goods in which they lack comparative advantage. Depending on the relative size of the countries in question and the demands for each good that they produce, the end result may be complete specialization (with no domestic production of certain goods) or partial specialization (simultaneous imports and domestic production of a particular good). The same principles apply to service sectors as long as the services are exportable (it is difficult to export a haircut).⁴

Of course, once trade opens, the autarky prices that motivate trade will change. In the example here, as guns flow into country B the price of guns relative to butter should fall, and vice versa in country A. The precise changes in prices that will result will depend on consumer demand in each nation, a complication that we need not introduce for present purposes. In *equilibrium*, the economic returns to engaging in the import-export business should be no greater than the returns to engaging in other activities. But trade will persist, for if it were to cease the price differences that gave rise to it in the first instance would resurface and trade would again yield especially high returns.

Notice also how little is necessary for a country to have comparative advantage in *something*. In our two-country, two-good illustration, *any* difference in the ratio of the price of guns to butter between the two countries in autarky ensures that one country has comparative advantage in one good and one in the other. Only if the price ratios were identical across the two countries would comparative advantage disappear.

Generalizing the analysis

The illustration in the section above employed several rather strong, simplifying assumptions. None are essential to the key analytical insights.

The assumption of two countries can readily be relaxed to allow for N countries, with N being large or small. The assumption of two goods can be relaxed to allow for M goods, with M being large or small. Comparative advantage is then determined by reference to a much more complex set of price ratios across nations, and the number of permutations in possible trade flows grows rapidly. But the basic notion of comparative advantage survives. Any difference in the autarky price ratios across nations for any two goods and any two countries implies a profitable trading opportunity, with a potential trade flow in the direction predicted as above. The harder question is

⁴ It is certainly possible, however, to export a barber. In the absence of personal or legal impediments to immigration, therefore, we would predict that skilled barbers will tend to migrate to nations where the price of haircuts in terms of other goods and services is high. Such movements of 'factors of production' – labour, capital, and the like — are a substitute for international trade in a most fundamental sense. Impediments to trade, therefore, result in heightened economic pressure for factor movements, and vice versa. I touch on this issue further in the section on 'Generalizing the analysis'.

whether an even more profitable trading opportunity exists with some third, fourth or *N*th country, with a different trade flow and pattern of specialization. Although the mathematics may quickly become daunting, the insight about potential trading profits survives.⁵

The assumption that labour is the only input into the production process is also inessential. The number of productive inputs can be arbitrarily large, and we can distinguish as finely as we like within categories of inputs (different types of skilled and unskilled labour, for example). All that matters to the analysis is that the autarky price ratios for the outputs of the production process be different across countries. A review of the discussion in the section above should make clear that it was this difference in price ratios that determined comparative advantage and created the profitable trading opportunities.

For the same reason, the assumption that markets are competitive in each country is inessential. Whatever the degree of competition or monopoly power in any market, a difference in price ratios leads to the opportunities and incentives suggested by the illustration.

The assumption of zero transportation costs was slightly more important to the illustration, but relaxing it does not change the key insights. In reality, transportation is never costless, and the question is whether the profit to be made from importing and exporting, transportation costs aside, is enough to cover the associated transportation costs and still leave a net profit. Sometimes the answer is no, in which case trade will not occur and autarky production and prices will persist. But where a profit net of transportation costs (and, by the way, other costs associated with customs inspections, tariff payments, and so forth) exists, trade will occur with specialization and exportation in accordance with the predictions of the theory of comparative advantage.

Finally, the assumption that labour does not migrate between country A and country B requires some discussion. Country A had absolute advantage by hypothesis, and in autarky its workers would have enjoyed a higher standard of living.⁶ If a worker in country B could move to country A and then

⁵ For a reasonably accessible discussion with references to the literature, see Ronald Jones and J. Peter Neary, 'Positive Theory of International Trade', in Ronald Jones and Peter Kenen (eds), *Handbook of International Economics*, Volume I (1984). A thorough treatment of some of the issues may be found in the same volume in Wilfred Ethier, 'Higher Dimensional Issues in Trade Theory'.

⁶ To see why, note that in autarky, a worker in country A endowed with one unit of labour per day could choose between consuming one gun or 1/2 unit of butter per day, or any linear combination of the two. A similarly endowed worker in country B could choose between 1/2 gun or 1/3 unit of butter per day, or any linear combination of the two. That is, the worker in country A would have a 'budget constraint' that lies entirely outside the budget constraint of the worker in country B, and could then enjoy more consumption of both goods and thus a higher standard of living.

become just as productive as the existing workers in country A, an incentive for them to move to country A would have existed. Concomitantly, if moving to country A gives workers from country B the same productivity as existing workers in country A, mass migration from B to A will increase global production relative to any conceivable allocation of resources without migration. A worker in country B specializing in butter production, for example, could produce still more butter per unit of time in country A because of country A's absolute advantage. Accordingly, the gains from migration in my illustration exceed any conceivable gains from trade, and mass migration will occur absent some impediment or cost to it that is outside of the 'model'.

More generally, if inputs into production (or 'factors of production') have greater productivity in one nation than in another, and if those factors are mobile, the result may be international factor movements rather than international trade in outputs. But just as the movement of goods is not costless, the movement of factors is not either. To take the case of labour, not only is the cost of transportation to another country potentially significant, but the costs of learning a new language, the costs of leaving behind culture and loved ones, and so on, can be great. Legal impediments to migration are also common. Thus, in a more general model, one must take all these costs and impediments into account for every factor of production. The equilibrium will entail some mixture of trade and factor flows. But because the costs and impediments to factor mobility are often considerable in practice, we can be confident that factor mobility is not a complete substitute for trade. At that point the theory of comparative advantage takes over to suggest how, in the face of limited factor mobility, international trade flows are likely to emerge.

Sources of comparative advantage

In the simple illustration above, comparative (as well as absolute) advantage arose because the productivity of labour was different across countries – a unit of labour produced different amounts of guns and butter in different locations. Put differently, we might say that the technology for the production of guns and butter was different in the two countries. Such differences across nations in the technology available for production are one source of comparative advantage, an idea usually attributed to the economist David Ricardo. Differences of technology may often be transitory, however, for the opening of trade is often accompanied by the exchange of information on other matters, including techniques of production.

Our illustration with only one factor of production (labour), however, masks another and often more important source of comparative advantage, identified in the work of Eli Heckscher and Bertil Ohlin. Some countries are rich in arable land; some have a large pool of well-educated workers; some have vast numbers of unskilled workers. Because of limited factor mobility,

such differences in ‘factor abundance’ will produce differences in the prices of various factors across countries. The Heckscher–Ohlin theory of comparative advantage suggests that nations will tend to have comparative advantage in the production of things that extensively use the factors of production that are relatively cheap and abundant locally. Thus, an abundance of arable land might create comparative advantage in wheat production; an abundance of skilled workers might create comparative advantage in the semiconductor industry; an abundance of unskilled workers might create comparative advantage in textile production. One can interpret this theory as subsuming the technology-based theory of Ricardo, since a nation with superior know-how in the production of a particular good can be characterized as being relatively abundant in the skilled factor of production (the skilled labour) that possesses such know-how.

A refinement in the theory that also blends the Ricardian and Heckscher–Ohlin perspectives is the ‘technology-gap’ idea generally attributed to Raymond Vernon. When a product is new and the technology to produce it is new, it will often require a fairly skilled labour force for its production. But as the technology becomes better understood, production becomes more routine and mass production with less-skilled workers will often become possible. Thus, nations with lots of innovators and skilled workers may tend to have comparative advantage in producing relatively new products, but over time comparative advantage may shift toward ‘imitator’ nations with less-skilled but cheaper work forces.⁷

Finally, some of the more recent work in trade theory suggests that comparative advantage may sometimes arise by chance. The basic idea, associated particularly with the work of Paul Krugman, is that certain production processes generate economies of scale and ‘learning-by-doing’ economies that reduce unit costs as the scale of production expands and as the quantity of output produced over time increases. Consequently, a nation that is a ‘first-mover’ in a new industry may capture comparative advantage simply by being first, as its costs fall more rapidly than those of later entrants.⁸

To some degree, nations may be able to affect their comparative advantage. Most obviously, a strong public education system will tend to produce comparative advantage in the production of things that require abundant skilled labour. Likewise, in the industries that fit the increasing returns pattern that Krugman emphasizes, nations may be able to generate comparative advantage by encouraging the establishment of new firms through tax incentives, subsidies or even protection against foreign competition for certain firms and industries. The harder question is whether and to what extent the nation (and the world as a whole) will benefit from such efforts. The

⁷ In addition to sources cited earlier, readers interested in these matters may wish to consult Edward Leamer, *Sources of International Comparative Advantage* (1984).

⁸ See generally Paul Krugman, *Rethinking International Trade* (1990).

net effects of protective trade policies, in particular, are usually thought by economists to be adverse. To those issues we now turn.

2. NORMATIVE ANALYSIS OF TRADE AND TRADE POLICY

The normative economic case for free trade, and the related ideas in this section, in one way or another rest on the premise that government intervention into international trade flows creates economic 'inefficiency', and that inefficiency is a bad thing. I thus begin with some background on welfare economics and the normative controversies surrounding the concept of efficiency before moving on to the case for free trade, its caveats, and various related issues in international trade policy.

Some concepts in welfare economics

Efforts to construct a 'social welfare function' with broad appeal to our shared instincts about policy are doomed to failure.⁹ Modern welfare economics thus abandons such efforts for the most part, and instead engages in some form of 'cost-benefit' analysis whereby the favourable and unfavourable effects of policy changes are compared to each other to determine whether the net impact of a change is positive or negative. Policy changes that produce a surplus of benefits over costs may be termed 'efficient', while policies with the reverse effect may be termed 'inefficient'.

A 'Pareto efficient' change in policy produces net benefits for at least one individual affected by it, and does not leave anyone worse off than before the change in policy. Such changes in policy seem compelling from a normative standpoint, as some people benefit on balance and no one is harmed. But it is exceedingly rare for a change in social policy to have this property. Accordingly, the welfare analysis of public policy usually inquires whether the benefits to those who gain from a change in policy exceed the costs to those who lose from it – if so, the change in policy is said to be 'Kaldor-Hicks efficient'.

To compare the net benefits of a change in policy for some individuals to the net costs of the change for others, it is necessary to convert those benefits and costs into common units. The natural unit for that task is money. Thus, when an economist states that a particular change in policy is 'efficient', that statement *usually* means that the monetary value of the benefits to those who gain from the change in policy exceed the monetary costs to those who are hurt by it.

1. Efficiency versus distribution

With this understanding of 'efficiency', it is immediately apparent that an 'efficient' change in policy may not be normatively appealing to everyone, and indeed one can readily imagine examples of efficient policy changes that

⁹ See Kenneth Arrow, *Social Choice and Individual Values* (1951).

seem quite unappealing to most of us. Consider, for example, a policy change that affects only two people – it will provide Donald Trump with an extra \$10, and it will take \$9.50 from a starving homeless person. This policy change is Kaldor-Hicks efficient, yet few of us would endorse it. One reason is that the ‘marginal utility’ of money is probably much greater for the starving homeless person than it is for Trump – that is, the happiness that \$9.50 can buy for the homeless person is potentially great and the happiness that an additional \$10 can buy for Trump seems minimal. Thus, even if the policy creates a net gain measured in money, we suspect that it on balance contributes to human misery. Such interpersonal comparisons of the ‘utility’ of money are difficult to make in most cases, but this case seems like a clear one.

As another example, consider a policy that takes \$950 away from an average working family, and gives 1¢ to each of 100,000 individuals chosen at random (\$1,000 in the aggregate). Again the policy is ‘efficient’, but because the costs fall very heavily on one family of moderate means, and the benefits are so dispersed as to seem trivial to each beneficiary, many of us might again find the policy unappealing.

The general lesson from these examples is that tensions that may arise between ‘efficiency’ (in the Kaldor-Hicks sense) and ‘distribution’. Most of us have instincts regarding a just distribution of income, and will feel that some efficient policy changes are nevertheless undesirable because of their effects on income distribution.

Does it follow from this observation that ‘efficiency’ is of little or no interest to us as a normative matter? One can argue ‘no’ for at least four reasons.

First, many changes in public policy do not have the perverse (to some) qualities of the above examples – a redistribution of income from poor to rich, or great burdens on the few that produce minimal benefits for the many. Some policies are indeed quite the opposite in their effects on the income distribution, and many others are approximately neutral in these respects. In the absence of evidence that a change in policy would affect the income distribution in a way that is troubling, one can argue that policies which increase the net monetary wealth of society are appealing. The reason is that greater monetary wealth, other things being equal, tends to increase human happiness because it can be converted into things that we value. In making this argument, of course, one reveals oneself to be a utilitarian, who solves the difficulties with interpersonal comparisons of utility across individuals by imagining that societal wealth is a good proxy for societal utility, at least where concerns about distribution can be put to one side.

Second, if *many* policy decisions are made on grounds of efficiency, each producing a net increase in societal wealth, it becomes much less likely that the income distribution will be affected in any important, adverse way. Instead, individual changes in policy will produce gainers and losers, but the aggregate effect of many policy changes that add to aggregate wealth may

well be to increase the wealth of everyone or virtually everyone. Put differently, if enough public policy decisions are made on grounds of Kaldor-Hicks efficiency, the collection of such decisions may produce a Pareto efficient change. Likewise, even if particular policy changes may have distributional consequences that trouble us, we might all agree 'behind a veil of ignorance' that our government should ignore such problems in the main and pursue efficiency because we each expect to come out ahead from efficient policies in the aggregate.

Third, even if we focus on a single efficient policy change and perceive that it has troubling distributional consequences, we must ask whether a further change in policy can remedy the distributional problems to our satisfaction without sacrificing the efficiency gains from the policy change under consideration. For example, suppose that a particular change in regulatory policy produces sizable efficiency gains in the Kaldor-Hicks sense, but that it systematically burdens certain lower income individuals while benefiting primarily upper income individuals. Before concluding that the policy change may be undesirable for that reason, we must at least ask whether an additional change in policy – such as a change in tax or welfare policy – might preserve the equity of the income distribution at lower cost to societal wealth than the perpetuation of the inefficient regulatory policy under scrutiny. More generally, if public policy decisions are to be made on grounds of distribution rather than efficiency, it is important to focus on the full array of policy instruments that are available for redistributing income. Some will carry greater costs to societal wealth than others, and it seems best to engage in 'efficient redistribution' whenever possible.

Finally, few would argue that societal wealth is irrelevant to public policy, even if other considerations are perceived as equally important or more important. At a minimum, therefore, welfare economics addresses a consideration that ought be factored into public policy decisions quite systematically. The efficiency of a change in public policy surely weighs in its favour, and might at least be said to afford a rebuttable presumption as to its normative appeal.

2. Global versus national welfare perspectives

To ascertain whether a change in policy is efficient, one must aggregate the effects on gainers and losers. Prior to that aggregation, one must specify the set of potential gainers and losers over which the aggregation is to be performed.

This task can become controversial when matters of international trade policy are under consideration. Suppose, for example, that a national government is considering whether to intervene in international trade in a particular way – whose welfare should 'count' in the policy-making process? Some of us may feel that the welfare consequences of national trade policies should be evaluated from the national perspective, with little or no weight

given to the welfare of persons outside the country, on the premise that governments ought to promote the welfare of their own citizens. But others may find this perspective troubling and prefer that national policy be set in a manner that is more other regarding.

Even for those who favour an emphasis on the welfare of their own nationals, the effects of trade policies on the welfare of others can become important as a practical matter. Policies that harm individuals in other countries may lead to retaliation, and thus indirectly harm domestic nationals in the end. Further, when nations begin to co-operate on matters of trade policy, as through the negotiation of trade agreements, the welfare of other nations enters matters directly as it will affect what their negotiators are likely to find agreeable.

Accordingly, the welfare economics of international trade can often usefully identify both the domestic and international welfare consequences of policy alternatives. The 'national' welfare perspective thus inquires whether a policy affords net benefits to domestic nationals. The 'global' welfare perspective looks at the costs and benefits to all affected parties wherever located, and gives all of them equal weight. Sometimes both perspectives yield the same conclusion regarding 'efficiency', but not always.

The simple case for free trade

We now turn to the welfare economics of trade policy, and in particular to the economic argument for free trade and its caveats. I will make this argument two ways – first by arguing that the opening of trade is efficient by comparison to autarky, and second by arguing that government restrictions on trade, as through a tariff or a quota in a particular market, are inefficient relative to letting the market achieve its own equilibrium.

To begin the discussion, recall the simple illustration above in which trade occurs between two countries, one with comparative advantage in gun production and the other with comparative advantage in butter production. Each nation specializes in the production of the good in which it has comparative advantage, and imports at least some of its consumption of the other good.

Such specialization of the labour force makes it possible for citizens in each country to consume a greater quantity of *both* goods. With reference to the input–output table above, suppose that citizens in both countries value both guns and butter, and suppose for purposes of illustration that in autarky, the labour force has been fully allocated in each country and that 100 units of guns and butter are produced in each country. Now imagine the following change in the allocation of labour: Country A shifts 20 units of labour from butter to gun production, sacrificing 10 units of butter production (it falls to 90) and increasing gun production by 20 units (it increases to 120). Country B shifts 36 units of labour from gun production to butter production, sacrificing 18 units of gun production (it falls to 82) and increasing

butter production by 12 units (it increases to 112). Country A then exports 19 units of gun production to country B in exchange for 11 units of butter production. Some simple arithmetic reveals that each country now has 101 units of both gun production and butter production to consume, more than they could possibly consume in autarky.

This illustration affords an important insight into the benefits of trade – specialization makes possible an increase in the global output of everything, holding constant the resources devoted to productive activity in each country, and thus makes it possible for trading partners to increase their consumption of everything without increasing the resources that they devote to production (only shifting them toward areas of comparative advantage). To be sure, trade may not in fact lead each nation to consume more of everything. The simple illustration here omits the demand side of the economy, and once consumer demands and prices are introduced it is assuredly possible that trade might lead nations to consume more of certain things and less of others. But if that happens, it is only because such a pattern of consumption is preferred.

If trade enhances the consumption opportunities of each nation relative to autarky, does it follow that trade makes every individual better off than in autarky? The answer, in general, is no. The removal of impediments to trade is likely to be Kaldor-Hicks efficient, but not in general Pareto efficient.

To appreciate the distributional impact of trade and the possibility that it may harm certain individuals, we must move away from our simple illustration in which a single factor of production (labour) produces everything and receives the same return regardless of where it is deployed. Suppose now that certain firms and workers are better at producing certain goods than others, and that they cannot shift to the production of other things without suffering a reduction in their profits or wages. This situation will arise at the firm level if, for example, the firm has invested in capital equipment that can be used to produce some things but not others. Similarly, at the worker level, workers may learn particular skills that are valuable in one industry but that do not transfer well to others (the knowledge of how to operate a loom, for example). In economic parlance, firms and workers have investments in ‘specific capital’ (physical or human) – investments that are specific to the industry in which they are working. Such investments are assuredly common in the economy. Further, certain factors of production such as arable land are by nature more valuable when devoted to certain activities rather than others, and their owners may thus be viewed as having specific investments as well.

Trade can harm domestic firms and workers who produce things that are in competition with imported things when those firms and workers have specific investments. An opening of trade in textiles, for example, will tend to lower the price of textiles because of the increase in supply from abroad. Firms and workers with specific investments in textile production will see a

decline in profits, wages or both, which cannot be avoided by moving into the production of other things because, by hypothesis, they cannot earn a comparable return on their specific capital investments elsewhere. Trade will also benefit these individuals by lowering the prices of other things that are imported, but the net impact on them may be adverse.

Trade may also have an adverse impact on certain consumers. The opening of foreign markets increases the demand for goods that are exported and thus raises their price. Other things being equal, this increase in price is detrimental to the domestic consumers who were previously able to buy those goods more cheaply. For many consumers the reduction in the price of imported goods relative to their price in autarky will more than compensate them, but this need not be true in all cases.

For these reasons, an increase in trade often creates both gainers and losers. The efficiency argument for trade must thus rest on Kaldor-Hicks efficiency. I will sketch this argument verbally, beginning with the import side.

When trade opens, the price of goods that are now imported will fall. Domestic consumers of those goods benefit while domestic firms and workers producing import-competing goods suffer. If domestic consumption and production remained exactly the same, the price gain to consumers, which measures the gain in consumer surplus per unit consumed, would exactly offset the price reduction to firms and workers, which measures the loss in producer surplus per unit. But neither consumption nor production remain as before. Consumers gain still more by increasing their consumption of the good in question, buying more units of the good to the extent that their willingness to pay for the good now exceeds the now lower price. Each such purchase yields additional consumer surplus. Likewise, firms and workers can cut their losses by shifting toward the production of other goods where the returns are greater. Hence, although the price of imported goods has fallen domestically, the attendant gain to domestic consumers is greater than the attendant loss to domestic firms and workers.

Putting the argument a bit more formally, let us suppose that, prior to trade, an economy consumes N units of some good at some price, say P . The firms and workers who produce this good domestically have specific investments. Trade opens and the price of this import-competing good falls by some amount, call it X . Now suppose, counterfactually, that domestic consumption of the good remains at N units, and that domestic production of the good remains at N units (nothing has changed but the price). Then, from a Kaldor-Hicks standpoint, the opening of trade would be neither efficient nor inefficient – consumers would pay a lower price [$P-X$ instead of P] for N units, and would thus come out ahead in dollar terms by an amount NX . But firms and workers would receive the lower price for their efforts, and would thus lose an amount NX . The two amounts cancel out and there would be no net gain or loss. But, of course, the opening of trade leads to more than just a drop in price. First, consumption will rise above N units

because as price falls, consumers who value the good at an amount less than P but more than $P-X$ will be led to purchase it. These consumers capture an increment in 'consumer surplus' equal to the difference between what they are willing to pay for a unit of the good and the price that they actually have to pay for. This increment in consumer surplus, aggregated across all consumers who are drawn into the market by the lower price, ensures that consumers in the aggregate will gain more than just NX – the amount NX is the gain to consumers who would have bought the good anyway at the higher price of P , while a further increment in consumer surplus accrues to those additional consumers who enter the market only because the price has fallen.

Second, as price falls, domestic firms and workers ordinarily will not continue to produce N units as before. Rather than bear the full wage/profit cut that would befall them if they continue to produce at that level, some will switch into the production of other things. They will do so at lower returns than before the drop in price because of their specific investments, but they are nevertheless better off doing so than by continuing to produce the same amount as before at the lower price. Consequently, firms and workers lose *less than* NX when price falls from P to $P-X$. They can reduce this loss to some extent by exiting the industry and doing something else.

Putting the two parts of the argument together, we have now shown that consumers in the aggregate gain more than NX when price falls from P to $P-X$, and that firms and workers lose less than NX when the price falls. This will be true for any drop in price, since X is completely arbitrary in the argument, and thus it will surely be true for the actual drop in price that results from the opening of trade. By the Kaldor-Hicks criterion, therefore, the opening of trade is efficient when we focus on the domestic consumers and producers of imports and their domestic substitutes.

An analogous argument establishes that gains arise on the export side. When trade opens, the price of goods that are now exported will rise. Domestic consumers of those goods are disadvantaged while domestic firms and workers producing export goods benefit. If domestic consumption and production remained exactly the same, the price detriment to consumers would exactly cancel out the price gain to firms and workers. But neither consumption nor production remain as before. Consumers cut their losses by reducing their consumption of the good in question, and shifting to other goods that yield them more surplus – such a shift will occur for every consumer whose valuation of the good in question is less than its price after the price increase. Likewise, firms and workers can take advantage of the higher price by increasing production of the good, drawing resources from the production of other things where the returns are lower. Hence, although the price of exported goods has risen domestically, the attendant loss to domestic consumers is smaller than the attendant gains to domestic firms and workers.

The opening of trade is thus Kaldor-Hicks efficient from the national

welfare perspective. And because this will be true for every nation that opens itself to trade, the opening of trade on a global basis must be Kaldor-Hicks efficient from the global welfare perspective as well.

If the opening of trade is beneficial to a nation, does it follow that any restriction on trade (any deviation from completely free trade) is harmful to that nation? The answer is ‘no’ in general, although under certain assumptions the answer becomes yes. In the remainder of this subsection, I will focus on the import side (since import restrictions are far more common in practice than export restrictions) and develop a scenario in which such restrictions are clearly inefficient. The next sections will then address some of the limitations and criticisms of this analysis. To keep the analysis as simple and accessible as possible, I will focus on a single market and discuss the effects of trade restrictions from a ‘partial equilibrium’ perspective that ignores their spillover effects into other markets via the impact on the price of foreign exchange. I assure the reader, however, that the argument still goes through in ‘general equilibrium’ under proper assumptions.¹⁰

Figure 1 depicts a competitive market for a single good in an importing nation. Since we are concerned with the question whether trade restrictions benefit or harm the importing nation, we employ the national welfare perspective and can ignore the effects of trade restrictions on outsiders. I will assume that the importing nation is ‘small’, so that it can purchase all of the good that it wishes from abroad at the ‘world price’ P . The supply curve for domestic production of the good is S , and the domestic demand curve for the good is D . With no intervention into the market by the government, total consumption would be determined by the intersection of the demand curve D with a horizontal line at the world price P (the horizontal distance Pf). Domestic firms would supply a quantity equal to Pe , and imports would supply the remaining consumer demand ef .

The diagram also gives us a depiction of the value of participating in the market to domestic consumers on the one hand and domestic firms and workers on the other. The value of participating in the market to consumers is equal to the difference between the most that consumers will pay for a good, and the price that they must actually pay for it. Aggregating across all consumers of the good, this amount is given by the area under the demand curve and above the price – the area of the triangle h/P – termed the ‘consumer surplus’.¹¹

In a competitive market, the supply curve reflects the marginal costs of production at each level of output. The gain to firms and workers from participating in the market is equal to the price of each unit sold, less its marginal costs of production. Aggregating across all units sold, the total gains

¹⁰ The texts cited in note 1 above provide general equilibrium treatments of varying complexity.

¹¹ Actually, this area is only an approximation to the consumers’ gain from participating in the market unless certain conditions are met. For a detailed discussion see Hal Varian, *Microeconomic Theory* (3d edn 1992), chapter 10.

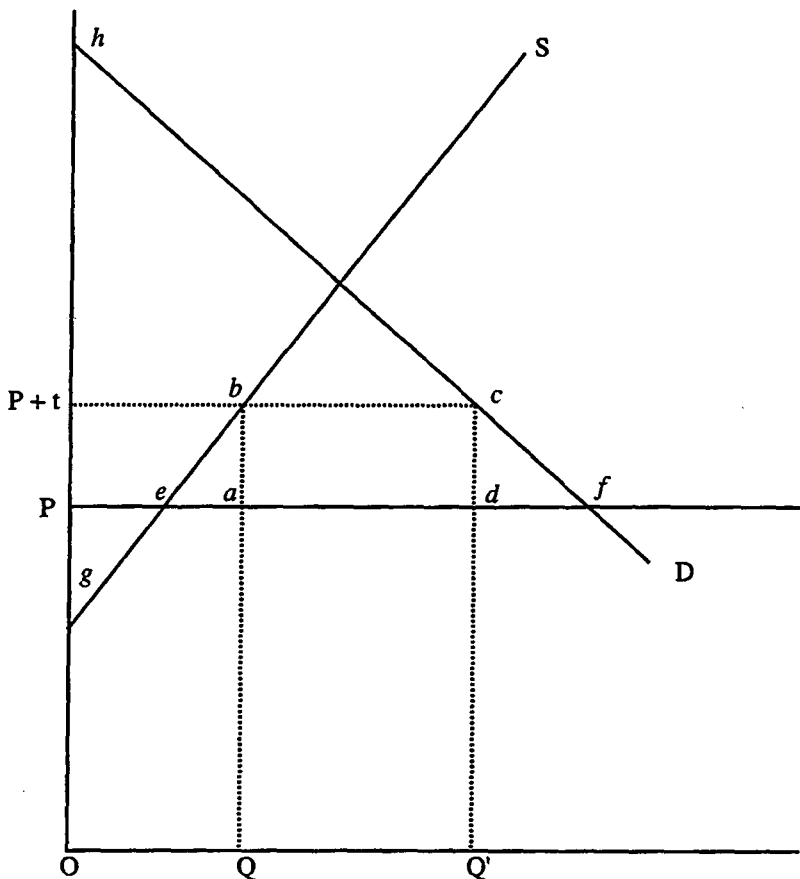


Fig. 1.

to firms and workers are given by the area above the supply curve and below the market price – the area of the triangle Pge – termed the ‘producer surplus’.¹²

Let us suppose, however, that the government decides to protect the domestic industry from foreign competition. It has various options for affording protection, including taxes on imports, known as ‘tariffs’, and restrictions on the quantity of imports, known as ‘quotas’. Consider first a tariff, and let the government impose a tariff equal to t on all units of the imported good. Because the importing nation is small and its foreign suppliers can sell elsewhere at the world price P , they will not reduce their price in response to the tariff, and thus the tariff will cause the price of the good to consumers to increase by the full amount of the tariff to $P+t$. Consumption will then equal Q' , with domestic production of Q and imports of $Q'-Q$. Domestic producer surplus rises to $(P+t)bg$. Consumer surplus has fallen to $(P+t)ch$,

¹² Ibid. at 224.

but government revenue on the $Q'-Q$ units of imports is now $(Q'-Q)t$, equal to the area *abcd*. Making the welfare comparison to the level of surplus that exists without government intervention into the market, the sum of consumer surplus, producer surplus and government revenue is now less by the two 'deadweight loss' triangles *eba* and *cdf*. Triangle *cdf* is associated with the loss of surplus to consumers who are priced out of the market; triangle *eba* reflects a further loss that is associated with using high marginal cost domestic resources to produce the good at home instead of buying it from abroad at the world price.

As an alternative to a tariff, the government could achieve the same degree of protection for the domestic industry by imposing a quota restricting imports to $Q'-Q$. Such a restriction of import supply would cause the price to rise until the market clears (total supply and demand balance). The market would clear when domestic production reached Q at the price $P+t$, for then domestic supply (Q) plus import supply ($Q'-Q$) would equal Q' , the quantity demanded at the price $P+t$. Such a quota would have the same impact on consumer and producer surplus as the tariff of t , but would not generate any revenue for the government (unless the rights to import under the quota were auctioned). Instead, the 'quota rents' *abcd*, equal to the tariff revenue under the tariff system, would be captured by the entities with the right to import under the quota. For example, if admission to the domestic market under the quota were offered to the foreign sellers that were first in line, those sellers would be able to command the elevated price $P+t$ for their merchandise and would receive an additional t per unit by comparison to the scenario without government intervention. The difference between the tariff and the quota in this setting, therefore, is not in the amount of surplus in the market, but in who captures it.¹³ From the national welfare perspective, the tariff and the equivalently protective quota have the same welfare effect if the quota rents are retained by domestic nationals, whereas the quota is far worse, other things being equal, if the quota rents are captured by foreign exporters.

For present purposes, however, the crucial point is that either a tariff or a quota is inefficient. The gains to domestic firms and workers in the industry in the form of increased domestic producer surplus, plus the gains to the treasury under a tariff or the gains (if any) to domestic nationals in the form of quota rents, will be smaller than the loss borne by domestic consumers in the form of reduced consumer surplus. This argument goes through for

¹³ This proposition oversimplifies but does no harm for present purposes. Important differences do exist between tariffs and quotas, for example, in imperfectly competitive markets. The notion that quotas are no worse than tariffs from a welfare standpoint in competitive markets is also questionable under many circumstances. For example, when quotas are allocated among supplier countries, they may not replicate the market shares for each nation that would obtain under a most-favoured-nation tariff system. Additional deadweight losses can then result because imported goods are no longer obtained from the most efficient suppliers.

any size tariff and any size quota. Under the assumptions used to construct Fig. 1, therefore, any restriction on trade in the form of a tax on imports or restriction on their quantity reduces national welfare.

Let us return briefly, however, to the issue of distribution. Conceivably, workers in the industry who would benefit from protection are relatively poor, and the consumers of the good in question are relatively affluent. Likewise, it is possible that protection confers large benefits on only a few domestic firms and workers of modest means, while the harm done to consumers is spread around among many and is quite small for each. Depending on the facts, therefore, it is conceivable that the effects of protection in this market, even if inefficient in the Kaldor-Hicks sense, may nevertheless affect the income distribution in a manner that might seem favourable to some observers, perhaps favourable enough to outweigh the efficiency loss from protection.

Before concluding that protection may be normatively desirable after all, however, it is important to ask whether other policy instruments may do a better job of enhancing the income distribution at lower cost to the economy. Welfare programs or changes in tax policy, for example, might provide a better method of aiding those who are perceived as deserving of it.

Indeed, economists conventionally argue that such tax and welfare measures are *far better* devices for redistributing wealth than protectionist measures like tariffs and quotas, for at least two connected reasons. First, trade protection is a crude way to aid those who are deserving of it. It raises prices in an industry, which will often benefit well-diversified, affluent shareholders in domestic firms as much or more than it will benefit truly deserving individuals (however we define them) who work in the industry. Tax and welfare policies, by contrast, can be targeted much more carefully so that the magnitude of the burden on the rest of the economy necessary to achieve any level of redistribution can be lessened.

Second, and closely related, properly designed tax and welfare policies entail fewer productive inefficiencies for a given degree of income redistribution than trade protection. Progressive income taxes and welfare subsidies for the deserving, in particular, create some inefficiency in the form of an adverse effect on work incentives. But consider the consequences of trying to achieve the same income distribution through protectionist measures. The adverse work incentive effects should be the same as a first-order approximation since, by hypothesis, the income distribution is the same either way. But protectionist measures entail the additional inefficiencies evident in the analysis above, namely, the deadweight loss associated with pricing some consumers out of the market and the deadweight loss associated with substituting higher cost domestic production for lower cost imports.¹⁴ It seems

¹⁴ This reasoning is just a variation on the general theme that inefficient regulatory measures are usually inferior to tax policies as a means of redistribution. See Steven Shavell, 'A Note on Efficiency vs. Distributional Equity in Legal Rulemaking: Should Distributional Equity Matter Given Optimal Income Taxation?', 71 *Am Econ Rev* 414 (Paps. & Procs. 1981).

exceedingly unlikely, therefore, that protectionist measures such as tariffs and quotas will prove the ‘cheapest’ way to aid individuals who are deemed deserving of public assistance.

Some misconceptions about free trade

The diagrammatic argument above embodies some explicit and implicit assumptions that may not hold in reality, and hence the claim that protectionist measures reduce national welfare requires some qualifications. Before considering these caveats to the argument for non-intervention, however, I wish briefly to address some arguments for protectionism that are often advanced but that are largely fallacious.

1. *Trade and jobs*

Ross Perot’s recent Presidential campaign in the USA devoted much energy to pressing the claim that trade liberalization – the North Atlantic Free Trade Agreement (NAFTA) in particular – would reduce the number of jobs in the American economy. Proponents of NAFTA responded with the claim that NAFTA would increase the number of jobs. This debate over the impact of trade on the number of jobs seems to resurface every time an important trade initiative arises.

To an economist, both sides of the debate ring hollow. Supply and demand in the labour market may be expected to balance in the long run just as it does in any other market, although transitory periods of disequilibrium may occur. The number of jobs in an economy will be determined by the labour supply and by the reservation wages of workers. A general decline in real wages may be expected to reduce the labour force somewhat as some workers exit to work in the home or to pursue leisure activities, and the reverse phenomenon may be expected if real wages rise. But absent a significant change in real wages economy wide, there is little reason to expect any significant change in the number of workers or jobs.

Incremental trade liberalization (such as that associated with NAFTA or the recent Uruguay Round) should not have much effect on real wages *on average*. To be sure, workers with specific human capital investments in import-competing industries can anticipate that their wages will fall with increased import competition, and the number of jobs in those industries will thus tend to decline. But workers in the export side of the economy will see increasing wages and employment opportunities as production shifts toward areas of greater comparative advantage. As a first order approximation, these effects may be expected to cancel out on average. Trade-related changes in the number of jobs, therefore, if any, are likely to be vastly smaller than other forces affecting job growth or loss such as the business cycle.

2. *Low wage competition*

Related to the argument that trade will reduce the number of jobs is the claim that trade between high-wage, developed countries and low-wage countries will tend to impoverish workers in the higher wage countries. In its popular form this argument is also flawed for much the same reasons as above – there is little reason to expect significant change in *average* real wages from the incremental trade liberalization that we have seen in recent years once the economy has fully adjusted to changes in trade patterns.

Further, if we shift the focus from wages to the overall standard of living, the effect of trade liberalization on the real *incomes* of citizens should be positive. The exploitation of comparative advantage enhances consumption possibilities for trading nations, as indicated earlier. The real purchasing power of citizens as a whole will rise, through an increase in the combined real returns to all factors of production.

These propositions do not rest on any assumptions about wage parity at the outset across trading nations. Low wages in developing countries should not be viewed as a threat to developed nations, but rather as a source of comparative advantage for developing nations that can usefully be exploited through trade.

To be sure, it is well known in the economics literature that an increase in the price of a good caused by trade will tend to increase the prices of the factors of production used intensively in the production of that good, while a decrease in the price of a good will tend to reduce the prices of factors of production used intensively in the production of that good (the Stolper-Samuelson theorem). Thus, workers whose skills are used intensively in the production of import-competing goods will tend to find their wages reduced by greater import competition. Further, if low-wage workers abroad are relatively unskilled and produce goods that tend to compete with goods produced by relatively low wage, unskilled domestic workers, then these comparatively less affluent domestic workers may bear the brunt of the effects of greater import competition. But this problem is really just a corollary of the distributional concerns noted previously, raising again the tension between efficiency and distribution. We then come full circle to the question whether protectionist trade policies are ever the best way to promote a more equitable income distribution. For the reasons given above, the answer seems likely to be no.

3. *National security and self sufficiency*

Proponents of protectionist policies also appeal with some regularity to concerns about national security. Industries from steel to textiles to mining to agriculture can be heard to argue that their presence and prosperity is essential to the ability of the nation to prosecute a war effort successfully.

To be sure, war does require armaments and munitions, the soldiers must have uniforms and they must somehow be fed. But it hardly follows that a

nation which may confront conditions of war must protect the industries that produce such items from decline in the face of foreign competition. Before arguments for protectionist policies based on national security can be taken seriously, a number of conditions must exist and a number of alternatives must be ruled out.

Many of the most important domestic productive facilities for national defence purposes will be supported in peacetime by government defence procurement – only where peacetime procurement is insufficient to support vital productive facilities do potential concerns about the contraction of domestic producers become important. Even then, the quantity of production of any particular good necessary to sustain a war effort and the domestic population during wartime may be considerably smaller than existing levels of domestic production. To take the case of steel in the USA, for example, the amount of production required for tanks, battleships and so on may be only a fraction of domestic capacity, and during large-scale conflict other uses for steel such as new automobiles and new building construction can be put on hold. Only where import dependence would still exist at those reduced levels of domestic consumption need the analysis proceed any further.

The likelihood of imports becoming unavailable in wartime must then be carefully considered. For a nation like the USA, serious interruption of seaborne commercial traffic seems unlikely to occur for most goods or commodities in any scenario short of global conventional conflict on the scale of World War II. The probability of such conflict seems small at best in the nuclear age. Further, in the event of an interruption in seaborne traffic, adjacent trading partners may be able to take up much of the slack on many items (in the case of the USA, Canada to the north and Mexico to the south).

Where interruption of necessary imports seems a serious risk, the next issue is whether domestic capacity can be restored with reasonable dispatch. Even if an industry has closed down certain productive facilities that might be needed in wartime, it does not follow that those facilities cannot be reopened or rebuilt quickly enough to satisfy essential needs.

Finally, stockpiling during peacetime may well be a superior alternative to the protection of domestic capacity. Where the item in question is not perishable, a nation might be better off by buying up a supply of vital material at low prices in an open trading system than to burden itself over time with the high prices attendant on protectionism as a hedge against armed conflict. The funds tied up in a stockpile have some opportunity cost to be sure; but this cost can easily be smaller than the costs of excluding efficient foreign suppliers from the domestic market. In the end, therefore, arguments for protectionism from the national security perspective require careful scrutiny and will rarely hold up to it.

Qualifications to the case for free trade

If some of the popular arguments against trade liberalization are dubious, economic theory has identified certain caveats to the argument for free trade that deserve mention. These caveats suggest that protectionist measures may under particular circumstances enhance national economic welfare, but usually at the expense of global welfare.

1. 'Optimal' tariffs

Figure 1 above was constructed on the assumption that the importing nation in question was 'small' and could purchase any quantity of imports that it desires at the world price. Concomitantly, it could not buy imports for an amount below the world price – in economic parlance, the importing nation was assumed to be a 'price taker'. As a consequence, any tariff that it imposed passed through in full to its domestic consumers.

In many instances, however, an importing nation may consume a large enough fraction of the locally available supply of a good to be able to affect the price that it must pay by varying the quantity that it purchases. More succinctly, the 'import supply curve' that the nation faces may slope upward, in contrast to the assumption in Fig. 1 that it is horizontal.

A tariff raises prices domestically, and the increase in prices to consumers causes them to reduce their purchases of the good in question. Thus imports must fall. If the supply curve of imports is upward sloping, a decline in the volume of imports will represent a movement down the supply curve to the left, and accordingly the unit price that the nation must pay for the imports that remain in the market will fall. The price to consumers in equilibrium will still equal the price of imports plus the tariff, as in Fig. 1, but because the price of imports has declined, the tariff does not pass through in full to domestic consumers. Rather, part of the tariff is 'paid' by exporters abroad who have reduced their prices in response to the tariff.

This possibility raises an important caveat to the welfare analysis in Fig. 1. The tariff there reduced national welfare because the decline in consumer surplus was greater than the increase in domestic producer surplus and government tariff revenue, producing the deadweight loss triangles discussed earlier. But where some of the tariff revenue is coming at the expense of foreign exporters – that is, where it represents an extraction of 'foreign producer surplus', the tariff need not reduce national welfare. Some deadweight loss will arise to be sure, but it is possible that the additional government revenue extracted from foreign sellers will outweigh the deadweight loss.

To give a simple example that may help convey the intuition, suppose that the import supply curve is highly inelastic – it slopes upward very steeply. This circumstance would arise if foreign sellers had few other outlets for their production (perhaps it is a perishable good that cannot be profitably exported to other countries and for which the domestic demand at home is

limited). Under these circumstances, foreign exporters would absorb most of any tariff by cutting prices. And because tariffs were mainly absorbed by foreign exporters, there would be little increase in price in the importing nation, little loss of consumer surplus, and little deadweight loss, yet potentially substantial government tariff revenue.

Indeed, the theory of the ‘optimal tariff’ imagines that nations with the power to influence the prices that they must pay for imports will set a tariff rate that maximizes national welfare (the sum of producer surplus, consumer surplus, and government revenue) in each market, taking account of the slope of the import supply curve for each good.¹⁵ The optimal tariff will be higher, other things being equal, the less elastic is the import supply curve (the more steeply it slopes upward).

Readers who are familiar with the microeconomic theory of the firm may notice a connection between this analysis and the theory of imperfect competition. A ‘monopsonist’ is defined in economics as an entity that has power over the prices of the things that it purchases (as distinguished from a monopolist which has power over the prices of the things that it sells). The monopsonist can increase its profits by reducing its purchases somewhat in order to induce its suppliers to lower prices. Optimal tariff theory is largely a translation of monopsony theory into the international trade arena, with nations setting tariffs to take advantage of their monopsony power over the price of imports.

Many trading nations are indeed large enough to have some power over the prices of their imports. If so, do the insights of optimal tariff theory seriously weaken the case against government interference with international trade? Most economists would respond ‘no’, for several reasons.

First, because the optimal tariff is an exercise of monopsony power, it introduces the same distortions into the international economy that monopsony does in a domestic setting. Price is elevated above the marginal costs of production for imported goods (by the amount of the tariff), causing deadweight losses for exactly the same reasons as in Fig. 1 (some consumers inefficiently exit the market and some high-cost domestic producers inefficiently enter it). Even though the tariff may be beneficial to the importing nation because of the extraction of surplus from foreigners, the effect on global welfare is clearly adverse. And if many nations engage in it, they may all discover in the end that their welfare reduced. Global co-operation to eliminate such behaviour is thus valuable.

Second, even if nations are unable to co-operate to eliminate the joint costs of optimal tariff policy, they must recognize that it is directly injurious to

¹⁵ Actually, they must also take account of the effect of their policies on other markets in ‘general equilibrium’, but in general equilibrium expositions of optimal tariff theory the source of gains from tariffs is the same. A classic reference is Harry Johnson, ‘Optimum Tariffs and Retaliation’, 21 *Rev Econ Studs* 152 (1952–53). J. Bhagwati and T. N. Srinivasan also give a nice treatment, from a general equilibrium perspective, in their Lectures (cited in note 1 above at 174–84).

trading partners because of the surplus that it extracts from them. Hence, nations cannot prudently assume that trading partners will react passively to optimal tariffs, and may instead anticipate that retaliation may be a common response.

Finally, the computation of the optimal tariff requires a tremendous amount of information about supply and demand conditions in each market. The 'general equilibrium' version of the theory suggests the need for even more information. It is unlikely that governments can reliably obtain this information, and even more unlikely that they can update it in timely fashion as market conditions change. Thus, efforts to impose the optimal tariff may well be doomed to failure for these reasons alone. Of course, the governments of large trading nations might impose a small tariff on the premise that it probably improves national welfare, retaliation aside, but the magnitude of any gain would be quite difficult to assess and the danger of overshooting the optimal tariff and causing harm would always be there. Not surprisingly, therefore, there is little evidence that government tariff policies have in practice been motivated by a search for the optimal tariff.

2. Externalities and distortions in other markets

A vast literature exists in economics on the problem of 'market failure' due to 'externalities'. Essentially, an externality is a benefit or a cost associated with the operation of market that is not reflected in the prices charged for a good or service or in the prices that purchasers are willing to pay for them. For example, automobiles emit pollutants that are harmful to society to some extent, but the seller and purchaser of the automobile bear only a tiny fraction of the harm caused by its emissions. Accordingly, in an unregulated automobile market, the 'social cost' of an additional automobile might be said to exceed its private cost to buyers because of the external cost to citizens who are not parties to the transaction. Automobiles would thus be priced too cheaply from an efficiency standpoint and society would tend to have too many of them around as a result.

Related to this is that markets may malfunction at least temporarily because of some impediment to normal market clearing processes. For example, John Maynard Keynes hypothesized that wages are 'sticky downward' and will not decline to balance supply and demand for labour during recessions (this notion has been severely criticized to be sure). If the sticky wage problem exists, then inefficient unemployment will result because workers are willing to work at a market clearing wage but are unable to do so.

Conceivably, international trade may contribute to such problems of market failure, at least indirectly. Imported automobiles may contribute to socially excessive pollution as much as domestic automobiles. Or import competition during a period of recession may exacerbate any inefficient unemployment that the economy is experiencing. Innumerable other examples might be offered.

Can the existence of such market failures justify restrictions on international trade? Without question, trade restrictions might at times enhance national welfare (and maybe even global welfare) relative to a policy of ignoring the market failure altogether. A tax on imported automobiles could raise automobile prices generally and reduce automobile consumption toward the efficient level. Restrictions on imports might result in more jobs for domestic workers in import-competing industries that were experiencing inefficient unemployment (although they might simply transfer the problem to exporting industries depending on circumstances).

Nevertheless, economists usually argue that such trade restrictions are inefficient because they are inferior to other forms of government intervention to correct the market failure.¹⁶ To take the automobile example, excessive automobile consumption is presumably not limited to imported vehicles. It may make sense for the government to reduce consumption by raising automobile prices, but it is best to do so with a tax on all automobiles, not just on imported ones. A tariff on imports may be able to induce the same target level of consumption if it is high enough, but it introduces a concurrent inefficiency in the production and consumption of automobiles that is unnecessary. In particular, it induces a substitution away from imports in favour of more costly or lower quality domestic vehicles, and creates an attendant deadweight loss. A non-discriminatory tax does not have that effect.

Similarly, in the labour market example, if the government knows that inefficient unemployment is present (not easy, to be sure), the cheapest way to remedy the problem is to intervene directly in the labour market to somehow eliminate the disequilibrium. Wage subsidies, mandatory wage reductions, and various other instruments can be used to bring the supply and demand for labour back into balance. And these instruments would not carry the distortion in the product market that results when the prices of imports are artificially inflated relative to the price of domestic substitutes.

3. The infant industry argument

A well-known argument for protectionism under certain circumstances, the 'infant industry argument', is really a variation on the market failure theme. This argument for protection rests on the notion that certain nascent industries take time to develop and lower their costs to the point of becoming competitive. With unfettered international competition, the argument runs, infant industries may be wiped out before they get going. But with a period of protection they can learn to operate profitably and ultimately survive and prosper in the international environment.

This argument has a number of difficulties with it. First, the ability of

¹⁶ Harry Johnson once again provides a classic reference: *Optimal Trade Intervention in the Presence of Domestic Distortions*, in *Trade, Growth and the Balance of Payments* (1965), chapter 11. J. Bhagwati and T. N. Srinivasan provide a succinct treatment in their *Lectures* (cited in note 1 above at 185–91).

governments to identify successfully those nascent industries that have the capacity to become competitive after a period of protection may certainly be questioned. Second, even where such an industry can be identified by the government, one must ask whether there is anything special about having it around domestically. Nations always have comparative advantage in some things, and even if they can create it in a new industry, what is the benefit? Only if the new industry would earn a higher rate of return than the resources that it consumes could earn elsewhere would special investments to create it be justified. Third, if the government can identify an industry with long-term potential to survive, and can determine that its rate of return is high enough to justify special investments to create it, why do the capital markets not finance its startup period without the need for government intervention? Before government measures in favour of an infant industry can be justified, therefore, we must have a very smart government coupled with a stupid capital market or at least a locally thin one that lacks the resources that are necessary.

But let us suppose that these objections can be overcome – what is the proper government policy instrument for nurturing the infant industry? The problem can again be viewed as one of market failure, this time in the domestic capital market. In essence, an investment with a positive social rate of return is not being financed because the cost of capital for some reason exceeds its opportunity cost (much like ‘sticky wages’ cause the price of labour to exceed the value of what workers can earn in their next best alternative). The ideal remedy, once again, would be an intervention directly into the capital market, such as a subsidy to commercial lenders to make startup loans, or even a government loan directly to the infant industry. Protectionist measures are inferior because of the deadweight losses that they create in the product market during the period of protection.

A possible answer to this argument might be that in some developing countries neither the local capital markets nor the government itself has enough liquidity to finance the startup period for the infant industry. But why not turn to foreign sources of capital at that point? Perhaps it might be argued that the government has trouble conveying the information that it has about the potential of the infant industry in credible fashion, so foreign lenders underestimate the likelihood of repayment. Or perhaps they are afraid of political risk and expropriation.

In the end, therefore, the infant industry argument cannot be altogether dismissed. But the number of conditions required to make it plausible in a given instance are extensive and unlikely to be met very often.

4. Strategic trade theory

Related to the classic infant industry argument for protection, though with some important additional wrinkles, are certain results in modern strategic trade theory. Recall the discussion of comparative advantage above, where it was suggested that comparative advantage may at times arise by chance

for ‘first movers’ in industries with increasing returns to scale (marginal costs that fall as output rises). Governments may be able to create comparative advantage in such industries by ensuring that their producers are the first movers or by stimulating domestic output to a level that exceeds what it would be without government intervention. Protection against import competition, thereby giving domestic firms a captive domestic market, is one way to promote more rapid growth in domestic output.

Once again, however, one must be careful to ask whether there are any gains from promoting the establishment of particular industries. If the firms and workers in those industries will in the end earn no greater return than they can earn elsewhere, and if domestic consumers will have access to the goods produced by the industry in question at a competitive price through imports, there is no reason to care whether the industry develops domestically or abroad.

Proponents of strategic trade policy might respond, however, that in industries with large economies of scale, it is often the case that very few firms survive in the long run because only a few are able to lower their costs enough to stay competitive. Hence, the producers that do survive will be oligopolists (or maybe even monopolists), earning returns that indeed exceed those available elsewhere in the economy. And if producers in a particular industry are destined to earn exceptional profits, governments might rationally prefer that those profits be captured domestically rather than abroad.

Another strand of strategic trade theory emphasizes not economies of scale to firms within an industry, but the possibility that firms in an industry produce positive externalities for the rest of the economy. For high technology industries in particular, it may be the case that it is difficult for firms to capture all of the returns to the innovations that they create. This possibility arises because of the various imperfections in patent, copyright and trade secret law, and the difficulty in writing and enforcing an ideal contract restricting employees from using a firm’s ‘intellectual property’ when they leave. Creative endeavour may thus generate positive spillovers, in the form of new knowledge, for other firms. It is possible that these spillovers do not cross national boundaries. For example, they may occur mainly as a result of employee turnover in an environment where international migration is quite limited.

If firms in an industry earn better than normal profits, or if they produce locally positive spillovers for other domestic firms, it is advantageous to a nation, other things being equal, to have them locate domestically. Does it follow that they should be protected from foreign competition to encourage their growth and expansion?

After the discussion in the last section, the reader will immediately wonder whether such protection is the best policy instrument for nurturing an economically ‘strategic’ industry. For familiar reasons, subsidies to the industry may be superior because they do not create the distortion in the product market that is created by protection. The taxes to finance the subsidies

create other distortions, of course, but this observation is not fatal to the argument that subsidies are probably superior. Of all the myriad ways to raise money for a subsidy, it would be the unusual case where every one of them creates a greater distortion than that created by protection against import competition.

But beyond the question of the proper policy instrument for nurturing an economically strategic industry lurk more fundamental difficulties. To exploit the insights of strategic trade theory, governments again require a great deal of information. Only some industries, probably quite a small percentage of them, fit the strategic trade scenarios. The ability of governments to identify them reliably, especially those fitting the positive spillover scenario, may be doubted.

In addition, where one government can identify a proper candidate for assistance, other governments probably can as well. The result may be a battle of subsidies or protectionist actions that are mutually offsetting and that cause substantial resources to be expended by governments for no national gain in the end.

Finally, efforts to exploit the insights of strategic trade theory, like the optimal tariff, tend to shift economic surplus around internationally without creating any more of it – that is, they may enhance national welfare but are likely detrimental to global welfare. A mutual agreement to eschew such behaviour would then be in the best interest of trading nations. In its absence, efforts to exploit the insights of strategic trade theory may simply invite mimicry and retaliation.

Discrimination in international trade

Nations can discriminate in their trade policies by restricting imports from some sources to a greater extent than from others. Typically, discrimination is accomplished through tariff rates that vary according to the national origin of the imports in question, but quotas and other protectionist devices can also create trade discrimination.

In general, trading nations that belong to the World Trade Organization (WTO) (virtually all major trading nations) are obliged to eschew discrimination in accordance with the ‘most favoured nation’ obligation. But this commitment has important exceptions, and discrimination is as a result quite common in practice. For example, developed nations regularly afford tariff preferences to developing nations. In addition, customs unions and free trade areas (such as the European Union and NAFTA) are permitted under WTO law and afford their members preferential access to each others’ markets. Since these arrangements all accomplish discrimination using tariffs, I will focus on tariff discrimination in the remainder of this section.

The economic effects of tariff discrimination on global welfare can be mixed. Discriminatory tariffs may enable relatively high cost producers in

the nations that benefit from lower tariffs to outcompete lower cost producers in the nations subject to higher tariffs. This phenomenon, known as ‘trade diversion’, creates deadweight losses that do not occur when all suppliers are subject to the same tariffs. Other things being equal, therefore, non-discriminatory tariffs enhance global welfare; a non-discriminatory tariff policy ensures that imports are supplied by the countries that can produce them most cheaply.

Yet, other things need not be equal, for it is readily possible that a reduction of tariffs on a discriminatory basis will improve welfare relative to the maintenance of higher, non-discriminatory tariffs. The reason is that any tariff reduction, discriminatory or otherwise, will allow lower cost foreign firms to displace higher cost domestic producers, saving resources in production and lowering prices to consumers – so-called ‘trade creation’. Thus, where the choice is between discriminatory trade liberalization and no liberalization at all, preferential liberalization can enhance global welfare.¹⁷ Furthermore, it is *possible* to construct a preferential arrangement so that such an improvement in welfare is guaranteed. The members of a new preferential arrangement could simply adjust the tariffs that apply to non-members so that trade with non-members remains the same as before. The welfare of non-members would then remain unaffected, and the members of the preferential arrangement would derive benefits solely from trade creation.¹⁸

In short, from a global welfare perspective, discriminatory arrangements are not ideal, and of course trade barriers of any sort are almost always incompatible with global welfare maximization. But discriminatory arrangements can be a second-best improvement relative to a more protectionist world without discrimination.

A similar mixed picture presents itself from the national welfare perspective. Trade preferences that favour a nation’s exporters will tend to be beneficial, and certainly can do no harm to national welfare. But when an importing nation grants trade preferences to the exports of others, it may injure itself if too much trade diversion results. For example, suppose that country A initially charges a uniform 20% tariff on widgets from all sources. Country B supplies all of country A’s market at a delivered price of 100 per widget, which translates into a price to the consumer of 120 after the tariff is collected. Now let country A grant a preference to widget producers in country C, and assume that widgets from country C enter country A duty free. Suppose further that widget producers in country C are considerably less efficient than in country B, and charge a delivered price of 119 per widget. Widgets from country C will sell at a price of 119 in country A’s

¹⁷ See Richard Lipsey, ‘The Theory of Customs Unions: A General Survey’ 70 *Econ J* 496 (1960).

¹⁸ See Murray Kemp and Henry Wan, ‘An Elementary Proposition Concerning the Formation of Customs Unions’ in Murray Kemp (ed.), *Three Topics in the Theory of International Trade: Distribution, Welfare and Uncertainty* (1976).

market, and will drive out widgets from country B because of their price advantage (assume all widgets are identical). The price to consumers in country A falls only slightly from 120 to 119, so the gain in consumer surplus per widget purchased is no more than 1.0. But for every widget that would have been imported from country B under the old tariff regime, the government has lost 20 in tariff revenue. The net impact on the national welfare of country A is likely adverse, by a wide margin, and the reason is that the tariff preference creates severe trade diversion – almost all of the tariff preference is wasted in the higher costs for producers in country C, which prevent consumer prices in country A from falling significantly despite the large tariff preference for country C. The government gives up a lot of tariff revenue and buys little for its consumers in return.

To conclude this section, I will briefly note one other wrinkle in the economics of trade discrimination from the national welfare perspective. The ‘optimal tariff’ for nations with monopsony power in international trade is discussed above. Implicitly, the discussion assumed that the tariff would be the same on all imports. But nations with monopsony power might do even more to extract surplus from foreign suppliers through tariff discrimination. The ideal form of ‘perfect price discrimination’ would involve a distinct tariff for each unit of each imported good equal to the difference between its supply price (given by its location on the import supply curve) and the competitive market price. Such a policy would extract all of the producer surplus from foreign suppliers without causing any distortion in consumption. It would not discriminate among countries but rather among producers, and perhaps even among different units of output from the same producer.

A loosely analogous form of imperfect price discrimination can be implemented such that goods from each country are charged a different tariff according to supply conditions in that country. For any price that prevails for consumers in the importing nation, an appropriately designed system of discrimination country by country can increase tariff revenue relative to what would otherwise be raised. In general, the revenue maximizing system of discriminatory taxes will tend to impose a greater tariff on exporting nations with a less elastic export supply curve, since producers in those nations will absorb more of the tariff through reduced prices than producers elsewhere.¹⁹

From the global welfare perspective, of course, such imperfect discrimination reduces welfare for much the same reasons as the optimal tariff. Likewise, as in the case of the optimal tariff, it may provoke retaliation. It is also quite difficult to implement, as the information necessary to practice it successfully will be difficult to obtain. Even if tariff discrimination by countries with monopsony power might in theory look tempting, therefore, it

¹⁹ For further discussion of the various possible motivations for trade discrimination, see Warren Schwartz and Alan Sykes, ‘Toward a Positive Theory of the Most Favoured Nation Obligation and its Exceptions in the WTO/GATT System’, 16 *Int'l Rev L & Econ* 27 (1996).

seems unlikely to be attractive in practice even absent international legal constraints against it.

A note on 'unfair practices': dumping and subsidies

Subject to the caveats and exceptions noted above, normative economic analysis argues against restrictions on imports from both the global and national welfare perspectives. Does the case for non-intervention change when imports are 'dumped' by foreign firms or 'subsidized' by foreign governments? A considerable literature exists on these questions, and it is useful before concluding this survey to touch on it.

Dumping occurs when a foreign firm engages in price discrimination *in favour* of a particular importing nation by offering customers there a better price than the price it charges to customers in the home market or in some third country market. Dumping can also occur if a firm charges prices that are below what is necessary to allow the firm to recoup all of its costs of production (including an allocation of its fixed costs) within a reasonable period of time. The first type of dumping – 'price discrimination dumping' – can be unwitting (perhaps a consequence of exchange rate fluctuations) or can occur because firms with a degree of market power often find it profitable to charge different prices in different markets. The second type of dumping – sales below 'cost' – often occurs during economic downturns when firms lower their prices because of weak demand.

Subsidization occurs when governments provide certain economic benefits to their producers. These subsidies often, though not always, result in a reduction in the price of goods produced by those producers (because the subsidy will often lower the marginal cost of production).

Why should an importing nation object to dumping or subsidization? Recall the discussion of Fig. 1 above, in which it was argued that protectionist measures such as tariffs reduce consumer surplus by more than they increase producer surplus and government revenue, thus causing an efficiency loss. Running the analysis in reverse, the same logic suggests that lowering the price of imports will increase consumer surplus by more than it reduces other components of national welfare, creating an efficiency gain. Nothing in that chain of reasoning in any way turns on the reason for the reduction in the price of imports – the fact that they are dumped or subsidized should make no difference. This observation has led most economic commentators to conclude that countermeasures against dumped or subsidized imports are impossible to justify from the national welfare perspective.²⁰ Rare cases in which dumping or subsidization might be associated with an attempt by a foreign government to monopolize world markets can perhaps be imagined (though

²⁰ See, for example, Janusz Ordober, Alan Sykes and Robert Willig, 'Unfair International Trade Practices', 15 *NYU Int'l L & Politics* 323 (1983); Alan O. Sykes, 'Countervailing Duty Law: An Economic Critique', 89 *Colum L Rev* 199 (1989).

I can think of none in practice to date), but a general hostility to dumped and subsidized imports seems inappropriate.

From the global perspective, the analysis is somewhat more subtle. Sales below 'cost' are normal in times of economic downturns and entirely unobjectionable unless associated with monopolization. Price discrimination dumping, by contrast, is a harder case where it is deliberate (accidental price discrimination is no concern). Like all forms of imperfect price discrimination, deliberate price discrimination across national markets has uncertain welfare implications. Relative to the behaviour of a non-discriminating firm with market power, a price discriminating firm does less damage in some markets and more damage in others, with the net effect being ambiguous.²¹ Of course, the damage is greater in the markets with the *higher* price (and thus the greater distortion due to the exercise of market power). Nations that get the lower price are coming out ahead relative to an environment in which no discrimination occurs.

Subsidization also presents some difficult issues from the global perspective. In principle, subsidies might be used to correct certain market failures, but they may also distort the market. A deadweight loss can occur from the global perspective if higher cost producers displace lower cost producers as a result of subsidies, or if goods are sold at a price below their true marginal cost of production as a result of subsidies. Once again, however, the nations that enjoy lower import prices because of foreign subsidies generally come out ahead, and any loss is borne by the nation that grants the subsidies in the form of a wasteful drain on the government treasury.

If countermeasures by importing nations against subsidized imports and against deliberate price discrimination dumping have the effect of discouraging such practices, therefore, it is conceivable (though by no means a certainty) that global welfare might improve. The irony is that the efficiency gains, if any, would be reaped by nations other than the nation taking the countermeasures.

CONCLUSION

The normative economics of international trade affords a powerful tool for the analysis of government policies toward international commerce. I have not attempted to survey every possible application of normative economics to international trade policy, but I have set forth the essential concepts in the hope that readers will now have the background to analyse and discuss particular issues through the economic lens.

Readers who are familiar with the details of trade policy in practice will no doubt have noted that governments often depart from what normative

²¹ See F. M. Scherer, *Industrial Market Structure and Economic Performance* (2nd edn, 1980), chapter 11.

economics might counsel. The reasons why are fascinating, and have spawned yet another strand of international economics research on the positive economics of trade policy formulation. I will survey the insights from that literature in a future contribution to this journal.