# On Subsidies in Trade Agreements*

Philip Sauré†
Swiss National Bank

This Draft: December 2008

**Abstract**

When terms of trade externalities are the only source of cross-border inefficiencies, trade negotiators must negotiate market access only. Based on this principle, Bagwell and Staiger (2006) argue against the WTO subsidy rules. The present paper shows that their argument fails when trade agreements are required to be self-enforceable. By affecting output, subsidies impact the self-enforcement constraints, which, in turn, determine the policies of trade agreements. Consequently, trade agreements must include subsidy rules. In realistic scenarios, banning production subsidies is efficient while negotiating market access only is inefficient. In this sense, this paper makes a strong case for the WTO subsidy rules.

Keywords: Trade Agreement, Subsidy, Self-enforceability.

JEL Classifications: F10, F13.

*I would like to thank Kyle Bagwell, Gino Gancia, Jaume Ventura and two anonymous referees for their helpful comments. All remaining errors are mine.

†Swiss National Bank, Börsenstrasse 15, 8048 Zurich, Switzerland. E-mail: Philip.Saure@snb.ch. The views expressed in this paper are the author's and do not necessarily represent those of the Swiss National Bank.

# 1 Introduction

Production subsidies have long plagued trade negotiations. Baldwin (2006) reports that "negotiations at Cancun collapsed [...] in the absence of greater commitments by the developed countries to reduce agricultural subsidies and lower import barriers on agricultural products." In fact, by blocking international cooperation, subsidies might induce economic losses that exceed even the heavy direct costs (see Anderson (2004)). Despite the obvious importance of subsidies for trade talks, trade theory has, until very recently, neglected their role in trade agreements. In a laudable attempt to fill this gap, Bagwell and Staiger (2006) provide the first formal analysis of the issue and show that efficient trade agreements target market access only and that, consequently, the *de facto* prohibition of production subsidies put in place by the current WTO legislation is inefficient. Their argument builds on the fact that all cross-border inefficiencies travel through world prices and that, moreover, world prices are uniquely determined by market access. Consequently, the authors conclude, negotiating mutual market access is sufficient to resolve all cross-border inefficiencies and any additional restriction – *e.g.*, a prohibition of subsidies – can only generate inefficiencies.

The present paper shows that this line of argument fails when trade agreements are required to be self-enforceable. By imposing self-enforceability the literature acknowledges the fact that sovereign countries cannot be forced into trade agreements but instead join and respect them only if that seems beneficial from each country's perspective individually. This assumption is truly restrictive, since, by the optimal tariff argument, large countries have incentives to defect from trade agreements by raising tariffs to distort world prices in their favour. Fortunately, however, countries refrain from doing so when the benefits of today's defection come at the cost of tomorrow's breakdown of international trade cooperation. Accordingly, the central precondition for trade agreements – formalized by the self-enforcement constraint – requires that the dynamic benefits of honoring an agreement outweigh those of defection.

Under this self-enforcement requirement, optimal trade agreements generally target their members' production subsidies for two fundamental reasons. First, the self-enforcement requirement does not constrain the individual country's policy choices but, instead, the policies a trade agreement can successfully implement. Hence, the self-enforcement constraint must be addressed by the designers of trade agreements – *e.g.*, by a supranational institution as the WTO – and not by the countries individually. Second, the value of cooperation and the value of defection, and thus the self-enforcement constraint, are affected by the output structures of member countries and thus by their subsidies. Consequently, the optimal trade agreement requires a specific amount of production subsidies. By the first observation the optimal level of subsidies needs to be specified by the optimal trade agreements.

In sum, the requirement that trade agreements be self enforcing makes a general case for addressing subsidies in their legal code. To exemplify this general principle and to evaluate the WTO subsidy rules in the light of self-enforceability, the present paper develops a two-country two-good competitive general equilibrium model, where benevolent governments employ import tariffs and production subsidies in order to maximize their citizens' welfare. Under the assumption that the self-enforcement constraints bind marginally, the model offers two new insights. First, under an optimal trade agreement production subsidies for import competing sectors are zero. Second, and as a corollary, trade agreements that target market access only are necessarily inefficient.

The intuition for the optimal ban of production subsidies is the following. As production capacities require time-to-build, output reacts to prices changes only with a delay. If a country uses subsidies repeatedly to create a continuous flow of import-competing output, it is less vulnerable to sudden import disruption than in absence of such intervention. Thereby, it mitigates the hardship of a breakdown of trade cooperation, which constitute the cost of defection and thus increases it's own defection temptation. This latter effect jeopardizes the efficiency of the trade agreement and constitutes the reason why subsidies are to be prohibited by optimal agreements.

By focussing on the role of production subsidies in trade agreements, the present paper closely connects to Bagwell and Staiger (2006). Very much in contrast to this earlier work, however, the present study provides a rationale for directly addressing subsidies in trade agreements and, in particular, for prohibiting them altogether. Thus, it may serve as a useful complement to Bagwell and Staiger (2006) in a balanced evaluation of the prevailing WTO legislation.

Since the early work of Yarbrough and Yarbrough (1986) and Dixit (1987), trade theory has generally understood trade agreements as a set of rules that encourage trade integration in the absence of a supra-national executive authority and which, therefore, must be self-enforcing. Prominent work include Devereux (1997), Maggi (1999), Bagwell and Staiger (2000), Park (2000), Ederington (2001), and Bond and Park (2002). Within this literature, some contributions highlight the effects of adjustment costs of output for optimal trade agreements. In presence of such rigidities, a change of regime from cooperative to non-cooperative policies is generally prolonged and more costly. In this case, the consequences of a defection on trade agreements are typically harsher, which, in turn, makes a defection less attractive. Via this channel adjustment costs can generate endogenous gradualism in trade liberalization when output changes sluggishly (Staiger (1994) and Furusawa and Lai (1999)). At the same time, adjustment costs induce endogenous shifts in outside options and bargaining positions, which, in turn, can generate an aggravated version of a hold-up problem. McLaren (1997) presents such a scenario and concludes that trade can make a "small country worse off than it would have been if its trade partner did not exist". Crucial to these results is the assumption that the supply side of the world economy is entirely decentralized. The present paper departs from this assumption to analyze how and when governments should intervene in decentralized production in order to alleviate efficiency losses. Moreover, it argues that in some cases the relevant subsidy rules must be fixed by the legal code of trade agreements.

Finally, Ederington (2001) argues that trade agreements need not target domestic policies (such as subsidies), showing that "trade policy is the most efficient means of countering the temptation to defect" on trade agreement and efficient trade agreements must use only tariffs to keep domestic policies at their individually optimal level. This statement, however only applies under a set of specific assumptions, which includes, in particular, the absence time-to-build requirements when production reacts to price changes instantaneously. The present paper, in contrast, precisely builds on the feature of sluggish production changes, which is both, arguably realistic and supported by empirical work (see Montgomery (1995) and Koeva (2000)).

The remainder of the paper is organized as follows. Section 2 introduces the general setup and formalizes the incentives to defect from trade agreements. Section 3 characterizes the optimal self-enforcing trade agreements and presents the main results of the paper. Section 4 concludes.

# 2 The Model

Based on a repeated trade game between two countries, the model developed in the following illustrates the role of production subsidies for international cooperation.

## 2.1 The Basic Setup

There are two countries, Home (no $^*$) and Foreign ($^*$), populated by individuals of masses $L$ and $L^*$, respectively. Individuals in both countries consume two final goods $X_1$ and $X_2$.

**Demand.** Consumers are infinitely-lived and derive lifetime utility

$$U^{(*)} = \sum_{t \geq 0} \beta^t \ u\left(c_{2,t}^{(*)}, c_{2,t}^{(*)}\right) \tag{1}$$

where $c_{i,t}^{(*)}$ is consumption of good $X_i$ at time $t$. The momentary utility $u$ (simply referred to as utility $u$) is continuously differentiable and gives rise to constant and equal expenditure shares on the goods $X_1$ and $X_2$. Aggregate consumption in either country is thus

$$c_1^{(*)} = I^{(*)}/2p_1^{(*)} \qquad c_2^{(*)} = I^{(*)}/2p_2^{(*)} \tag{2}$$

where $I^{(*)}$ are national incomes and $p_i^{(*)}$ are local prices. Here and whenever there is no risk of confusion, time indices are omitted.

**Supply.** Competitive firms in both countries produce the goods $X_i$ using a constant returns to scale technology with labor and sector-specific intermediate goods $Z_i$ as the factors:

$$x_i^{(*)} = a_i^{(*)} f_i(z_i^{(*)}/L_i^{(*)})L_i^{(*)}. \tag{3}$$

where $L_i$ and $L_i^*$ stand for the countries' labor allocation to sector $i$ while $z_i$ and $z_i^*$ represent the quantities of the intermediate goods $Z_i$. The functions $f_i$ are twice continuously differentiable and satisfy $f_i(0) = 0$, $f_i' > 0$, and $f_i'' < 0$.

Home is assumed to have a comparative advantage in the $X_1$-sector, reflected by

$$\frac{a_1}{a_2} > \frac{a_1^*}{a_2^*}. \tag{4}$$

Intermediate goods $Z_i$ are produced with the constant returns to scale technology

$$z_i^{(*)} = L_{z_i}^{(*)}. \tag{5}$$

Unless otherwise noted, markets are competitive in both countries. The technologies (3) and (5) imply that labor is the only true production factor and the model is essentially Ricardian. Assumption (4) introduces motives to trade.

**Integrated Economy.** Intermediate goods $Z_i$ are produced competitively through (5) so that the price of the intermediates equals the wages. The intermediate-to-labor ratio in production of final goods will be denoted by $\xi_i = z_i/L_i$. Cost minimization of firms in the final good sectors implies[1]

$$1 = f(\xi_i)/f_i'(\xi_i) - \xi_i \tag{6}$$

as long as output in the relevant sector is positive. These conditions determine the ratios $\xi_i$. By $f_i'' < 0$ and $f_i > 0$ the expressions $f_i(\xi_i)/f_i'(\xi_i) - \xi_i$ are increasing in $\xi_i$ and the ratios $\xi_i$ are unique.

In the following $X_1$ is chosen as the numeraire so that the price of $X_2$ coincides with the relative price between $X_2$ and $X_1$, denoted by $p$. With this convention, a competitive labor market, together with technologies (3) and (5), implies $a_1 f_1'(\xi_1) = p a_2 f_2'(\xi_2)$ so that the equilibrium price is unique:

$$p = \frac{p_2}{p_1} = \frac{a_1 f_1'(\xi_1)}{a_2 f_2'(\xi_2)} \tag{7}$$

Together with aggregate income $I = wL = a_1 f_1'(\xi_1)L$, price (7) determines aggregate consumption $c_i$ by equations (2). Equilibrium labor allocation in the $X_i$-sector (including labor employed in intermediate production) equals

$$(1+\xi_1)L_1 = L/2 \qquad \text{and} \qquad (1+\xi_2)L_2 = L/2. \tag{8}$$

By the first fundamental welfare theorem, the equilibrium thus defined is efficient.

**Free Trade.** When final goods are costlessly tradable Home and Foreign specialize according to their comparative advantages. The aggregate production functions in each country can be summarized by the technologies

$$x_i = A_i L_i \qquad \text{and} \qquad x_i^* = A_i^* L_i^* \tag{9}$$

where

$$A_i^{(*)} = a_i^{(*)} \frac{f_i(\xi_i)}{1+\xi_i}. \tag{10}$$

Notice that the $\xi_i$ are determined by (6) and do not differ across countries or trade regimes, which implies that the comparative advantages are still determined by condition (4), implying $A_i/A_i^* = a_i/a_i^*$ and $A_1^{(*)}/A_2^{(*)} = a_1^{(*)}/a_2^{(*)}$. In equilibrium the value of world expenditure on $X_1$

[1] Solve min $_{L_i,\xi_i} L_i(1+\xi_i)$ s.t. $a_i f(\xi_i) L_i \geq 1$.

equals the value of supply of $X_1$. Under constant expenditure and equal shares this condition is $p_1(x_1 + x_1^*) = (p_1(x_1 + x_1^*) + p_2(x_2 + x_2^*))/2$ and determines the relative price of $X_2$ to $X_1$ as

$$p = \frac{p_2}{p_1} = \frac{x_1 + x_1^*}{x_2 + x_2^*}. \tag{11}$$

At the same time cost minimization of firms implies

$$p \in \begin{cases} \frac{a_1}{a_2} & if\ x_1, x_2 > 0 \\ \left[\frac{a_1}{a_2}, \frac{a_1^*}{a_2^*}\right] & if\ x_2 = x_1^* = 0 \\ \frac{a_1^*}{a_2^*} & if\ x_1^*, x_2^* > 0. \end{cases} \tag{12}$$

Equations (11) and (12) determine the unique world price $p$ of the respective regime. Labor allocation and output are given by (3), (6), and $\xi_i = z_i/L_i$ and the technologies (9). Finally, national incomes $I^{(*)} = (x_1^{(*)} + px_2^{(*)})$ uniquely determine consumption through (2).

As in standard Ricardian models, three possible patterns of international specialization emerge. First, partial specialization with Home producing $X_1$ and $X_2$ and Foreign producing $X_2$ only; second, full international specialization with Home producing $X_1$ only and Foreign producing $X_2$ only; and third, partial specialization with Home producing $X_1$ only and Foreign producing $X_1$ and $X_2$. Notice that at least one of the countries completely specializes in production of one of the goods. Throughout the paper the comparative advantage is assumed to be strong enough to generate full specialization under free trade. This amounts to assuming

$$1 < \frac{a_1 L}{a_1^* L^*} \qquad \text{and} \qquad 1 < \frac{a_2^* L^*}{a_2 L} \tag{13}$$

This limitation is less demanding than it seems at first sight. As discussed below, the key requirement for the paper's findings is that each country produces world supply of at least one good under free trade. This requirement on international specialization emerges in a great variety of standard trade models (see e.g. Dornbusch et al (1977), Dornbusch et al (1980), Acemoglu and Ventura (2002), and Romalis (2004)) and is, in this sense, an entirely natural one.

## 2.2 Government Policies

Governments are assumed to set production subsidies and import tariffs so as to maximize their citizens' utility. The definition of both policy instruments will be given next.

**Import Tariffs.** Governments in Home and Foreign can set gross ad valorem import tariffs $T$ and $T^*$. Throughout the paper Home's domestic price of good $X_1$ is taken as the numeraire and the world price of $X_1$ denoted with $p$. Thus, domestic prices in Home and Foreign for goods $X_1$ and $X_2$ are

$$p_1^{Home} = 1 \quad p_2^{Home} = Tp \quad \text{and} \quad p_1^{Foreign} = T^* \quad p_2^{Foreign} = p.$$

Home's national income includes tariff revenues and equals $I = x_1 + Tpx_2 + (T-1)p(c_2 - x_2)$ so that with $c_2 = I/(2Tp)$ from (2)

$$I = 2\frac{T(x_1 + px_2)}{T+1}. \tag{14}$$

Foreign's income is $I^* = T^*x_1^* + px_2^* + (T^*-1)(c_1^* - x_1^*)$ or

$$I^* = 2\frac{T^*(x_1^* + px_2^*)}{T^*+1}. \tag{15}$$

The trade balance $p\,(c_2 - x_2) = c_1^* - x_1^*$ together with (2), (14), and (15) determines relative prices

$$p = \frac{x_1(T^*+1) + x_1^*T^*(T+1)}{x_2T(T^*+1) + x_2^*(T+1)}. \tag{16}$$

**Production Subsidies.** Recognizing the myriad ways to subsidize production, the World Trade Organization gives a very broad definition of subsidies. Article 1 of the Agreement on Subsidies and Countervailing Measures defines a subsidy as "a financial contribution by a government or any public body," where financial contributions can consist of "a direct transfer of funds," "revenue that is otherwise due... [but] not collected" or the provision of "goods or services other than general infrastructure" (see WTO (1995)). The present paper adopts the last version as a definition of a subsidy. Following the WTO code, a subsidy will be defined as a public provision of the sector-specific input $Z_i$. More precisely, to subsidize sector $X_i$, a government purchases the amount $\bar{z}_i \geq 0$ of the intermediate good $Z_i$ from one or more private firms in the intermediate sector. This purchase is realized through a price-guarantee the government gives the $Z_i$-producers for the pre-determined quantity $\bar{z}_i$. The guaranteed price must weakly exceed the firms' production costs $w$. The government-controlled quantity of the intermediate goods, $\bar{z}_i$, is distributed to a set of final good firms without charge; the costs are financed through lump-sum taxes.[2]

Since government activities are usually thought to be inefficient, a fraction of the intermediate

[2] The price-guarantee for $Z_i$-producer as well as the distribution of the $\bar{z}_i$ may generate positive profits. Under homothetic preferences total income is the only determinant for aggregate demand and the specific distribution of profits is irrelevant.

good is assumed to be lost in the process of subsidization and the effective unit labor requirement is $\gamma > 1$ for production of the subsidized intermediate goods $Z_i$. Throughout the analysis, subsidies will be given to import-competing sectors only, i.e. $\bar{z}_1 = \bar{z}_2^* = 0$.[3]

Figure 1 illustrates the impact of a subsidy $\bar{z}_2 > 0$ in Home on its production possibility frontier. There are two main consequences of the subsidy. First, the production possibilities are strictly reduced due to the inefficiency in centralized production ($\gamma > 1$). Second, the subsidy implies that production of good $X_2$ is positive at low relative prices $p^{Home}$ when it would be zero in absence of subsidies. In particular, the condition $\lim_{x\to\infty} f_i(x) = \infty$ implies that domestic production of $X_i$ is positive at all finite prices $p^{Home} \in (0, \infty)$, whenever $\bar{z}_2 > 0$.

**Production under Subsidies and Tariffs.** Under positive subsidies the intermediate labor ratio in Home's sector $X_2$ is no longer determined by (6). In general, profit maximization in the competitive domestic market implies

$$Tpa_2 \left[f_2(\bar{\xi}_2) - \bar{\xi}_2 f_2'(\bar{\xi}_2)\right] = w \tag{17}$$

for the distorted ratio $\bar{\xi}_2$. This condition implies $0 < \bar{\xi}_2 < \infty$ for finite real prices $Tp/w \in (0, \infty)$. Thus, governments can generate positive import-competing production by subsidizing the according sectors. More precisely, one can show the following claim.

**Claim 1** *Assume $\bar{z}_2 = 0$ implies $x_2 = 0$. Then,*

*(i) $x_2$ is continuous and increasing in $\bar{z}_2$ around $\bar{z}_2 = 0$ and*

*(ii) average productivity of $X_2$-production is positive and finite.*

*The equivalent statement holds for $X_1$-production in Foreign.*

**Proof.** By assumption $\bar{z}_2 = 0 \Rightarrow x_2 = 0$ and expression (10) $a_2 f_2(\bar{\xi}_2)/(1+\bar{\xi}_2) < w$ holds. Consider now $\bar{z}_2 > 0$ and assume that the equilibrium $z_2$ satisfies $z_2 > \bar{z}_2$. This implies $Tpa_2 f_2'(\bar{\xi}_2) = w$ and, with (17), $Tpa_2 f_2(\bar{\xi}_2) = w(1 + \bar{\xi}_2)$, thus contradicting $a_2 f_2(\bar{\xi}_2)/(1 + \bar{\xi}_2) < w$. Consequently $\bar{\xi}_2 = \bar{z}_2 L_2$ holds.

*(i)* By (3), (5), and (17) $Tpa_2 \left[f_2(\bar{\xi}_2) - \bar{\xi}_2 f_2'(\bar{\xi}_2)\right] = f_1'(\xi_1)$ holds. As $\xi_1$ is constant by (6), this identity implies that $\bar{\xi}_2$ is decreasing in $p$. Further, (3) and concavity of $f_2$ imply that $X_2$-output

[3] This is a convenient but innocuous simplification, since the present paper's main mechanism is exclusively driven by the creation of import-competing production. It may also reflect the fact that state intervention tends to support import-competing industries.

$x_2 = a_2 \bar{z}_2 f_2(\bar{\xi}_2)/\bar{\xi}_2$ is decreasing in $\bar{\xi}_2$. Consequently, for given $\bar{z}_2$, output $x_2$ is increasing in $p$. As the world price (16) decreases in $x_2$ (and increases in $x_1$), given $\bar{z}_2 > 0$ the price $p$ and ratio $\bar{\xi}_2$ are uniquely determined and, moreover, continuous in $\bar{z}_2$. For continuity at $\bar{z}_2 = 0$, observe finally that by (17) $\bar{\xi}_2$ approaches a positive constant at $\bar{z}_2 \to 0$. Thus, $\lim_{\bar{z}_2 \to 0} x_2 = 0$ holds and $x_2 > 0$ for $\bar{z}_2 > 0$. This proves the claim.

*(ii)* Since output is $a_2 f_2(\bar{\xi}_2) L_2$ and labor input is $\left(1 + \gamma\bar{\xi}_2\right) L_2$, total labor productivity is

$$\bar{A}_2 = \frac{a_2 f_2(\bar{\xi}_2)}{1 + \gamma\bar{\xi}_2} \tag{18}$$

which is positive and finite as $0 < \bar{\xi}_2 < \infty$ by (17). ■

The first part of the claim shows that countries can use subsidies to generate positive production in sectors that would be idle without government intervention. According to the second part, the average productivity of induced $X_2$-production is finite and hence the efficiency losses are at most proportional to import competing production.

In sum, both countries can induce import competing production at the marginal rates of substitution of, respectively,

$$A = \frac{\bar{A}_2}{A_1} = \frac{1+\xi_1}{a_1 f_1(\xi_1)} \frac{a_2 f_2(\bar{\xi}_2)}{1+\gamma\bar{\xi}_2} \qquad \text{and} \qquad A^* = \frac{\bar{A}_1^*}{A_2^*} = \frac{1+\xi_2}{a_2 f_2(\xi_2)} \frac{a_1^* f_1(\bar{\xi}_1^*)}{1+\gamma\bar{\xi}_1^*} \tag{19}$$

Condition (12) on the world price $p$ then becomes

$$p \in \begin{cases} \frac{A_1}{T\bar{A}_2} & if\ x_1, x_2 > 0 \\ \left[\frac{A_1}{T\bar{A}_2}, \frac{T^*\bar{A}_1^*}{A_2^*}\right] & if\ x_2 = x_1^* = 0 \\ \frac{T^*\bar{A}_1^*}{A_2^*} & if\ x_1^*, x_2^* > 0. \end{cases} \tag{20}$$

Equation (17) implies that under positive domestic prices of the import good ($0 < Tp < \infty$), the ratio $\bar{\xi}_2$ (ratio $\bar{\xi}_1^*$) is strictly positive. Thus, the marginal rates of transformation $A^{(*)}$ defined in (19) are positive and finite.

# 3 Trade Agreements

It is well known that efficient trade agreements may implement non-zero tariffs, as long as relative prices are undistorted (see Mayer (1981) or Dixit (1987)). Consequently, the first best trade agreements are characterized by $p_1^{Home}/p_2^{Home} = p_1^{Foreign}/p_2^{Foreign}$ or $TT^* = 1$. In this case tariffs simply serve to transfer income from one country to another, without distorting neither supply nor demand.[4] The actual amount and the direction of these transfers depend on the countries' gains from trade, their outside options, and respective bargaining power.

The present paper, however, starts from the premises that the first best trade agreement is not self-enforceable but countries would, instead, defect on this unconstrained optimum. Fundamental to the concept of self-enforceable trade agreements is the optimal tariff argument, according to which large countries benefit from unilaterally charging tariffs, thus distorting the terms of trade to their favor. Put differently, countries have incentives to defect, or cheat, on a trade agreement by charging tariffs unilaterally. Such a defection, however, is typically assumed to come at the cost of a breakdown of future cooperation. The self-enforcement requirement is met, if, in a dynamic sense, the value of respecting a trade agreement be higher than the value of defecting on it.

### 3.1 Self Enforcement Constraint

Under a self enforceable trade agreement both countries receive weakly more utility from respecting the agreement than from defecting on it. Formally, let $\bar{u}^{(*)}$ represent the utility under cooperation, $u^D$ ($u^{D,*}$) under one-sided defection of Home (Foreign), and $u^{N,(*)}$ under uncooperative (or Nash) policies on both sides. With this notation and with the total utility (1) the self-enforcement requirement can be written as[5]

$$\sum_{s \geq t} \beta^{s-t} \bar{u}_s^{(*)} \geq u_t^{D,(*)} + \sum_{s \geq t+1} \beta^{s-t} u_{s,t}^{N,(*)} \qquad \forall \;\; t \geq 0 \tag{21}$$

This sequence of constraints collapses to just two constraints – one for Home and one for Foreign – when all parameters and equilibrium policies are time-invariant. The following analysis will be restricted to this case of time-invariant trade agreements.[6] Hence, the self-enforcement constraint in period $t = 0$ stands for the self-enforcement constraint of all periods.

The impact of tariffs and subsidies on the different components of (21) is central to the following analysis. The net gains from defection, and defection utility $u^{D,(*)}$ in particular, crucially depend on the timing of actions, which will be specified in turn.

[4] Such transfers can of course be realized through side-payments without involving negative net tariffs.

[5] Following the literature, I assume that countries act cooperatively in all periods after defection. Notice that the punishment utility $u_{t,s}^{N,(*)}$ may depend on the actual period $s$, and on the period of defection $t$.

[6] In terms of efficiency, this restriction is not innocent, as shown by the work of McLaren (1997), Park (2000), and Bond and Park (2002). It may be justified by restrictions on simplicity of legal codes that negotiators face.

**The Timing.** The literature's standard interpretation of defection is a unilateral deviation to the individually optimal policy in one period – the defection period – by one of the countries. This deviation is an off-equilibrium action and is thus unanticipated by other players. Hence, no players (other than the defector) can react within the same defection period but adapt their strategies in the following period. Consequently, the tariffs of the country which is defected on remain at the levels fixed by the agreement. The same holds true for private firms, which cannot react within the period of defection but need time to adapt their output.[7] Since this principle applies to firms in both countries, the world output structure, summarized by the vector

$$X = (x_1, x_2, x_1^*, x_2^*) \in \mathbb{R}_+^4$$

is fixed and taken as given within the defection period. Production subsidies, in the sense specified above, do not resolve this rigidity, since they, too, operate via price incentives for private firms.

A strictly positive reaction time of the economic agents that face with the unanticipated defection is an essential assumption of the repeated trade game. Indeed, this reaction time defines the length of the game's periods. In standard trade models, where the import tariffs are the only policies that governments set, this definition is unambiguous and clear. In the present model, however, where both import tariffs and output adapt to price changes, the central question arises whether tariffs and outputs are equally fast to change. Notably, the time a country needs to adapt its import tariffs may generally differ from the time required to bring into being an new industry in the country. On the micro level, firms tend to have extended adjustment periods to realize substantial increases in output plans. Capacity-building or start-up periods substantially limit private firms' short-term expansion of production when prices increase. Not surprisingly, empirical literature finds that adjustment of output is sluggish. Thus, Montgomery (1995) estimates that firms' "construction periods average five to six quarters" and Koeva (2000) estimates "that the average construction lead time for new plants is around two years in most industries." Compared to time spans of these dimensions tariffs are set and adapted rather quickly.[8]

These facts are built into the model through the assumption that governments can change the tariffs on a period by period basis while firms need $M > 1$ periods to increase output capacity. This assumption has important implications for the timing of events in case of a defection on trade agreements: after an unanticipated defection firms in both countries need $M$ periods to expand

[7] This is a crucial difference to the setup in Devereux (1997).

[8] Even under a lengthy WTO dispute settlement process the standard procedure takes about a year (see WTO (2007)). A clear-cut defection in the game-theoretic sense is likely to generate much quicker reactions.

production. Consequently, following a defection in period $t_0$, national output cannot expand for the periods $t \in \{t_0, t_0+1, .., t_0+M\}$ in both countries. From period $t_0+M+1$ onward, the output structures in both economies adapt according to production incentive (i.e., prices and subsidies). In sum, a defection is followed by two qualitatively different punishment phases: the first punishment phase – representing the medium term – during which uncooperative behavior is limited to tariffs while output is still at its cooperation levels and the second phase – representing the long term – characterized by uncooperative tariffs and subsidies with private firms producing accordingly.

With these assumptions about the timing, the self-enforcement constraints (21) take a particular structure, which will be formalized with the following notation. First, let $\bar{T}^{(*)}$ denote the tariffs set by the trade agreement and $\bar{X}$ the equilibrium world output structure under tariffs and subsidies of the trade agreement. Further, write $T^{BR}(T^*, X)$ and $T^{BR,*}(T, X)$ for Home's and Foreign's unilaterally optimal tariff given the respective other country's tariffs and given the world output structure. Moreover, let $T^{N,(*)}(X)$ be the tariffs of the Nash Equilibrium of the tariff game at given output $X$ (i.e. $T^N(X) = T^{BR}(T^{BR,*}(T, X), X)$ and $T^{N,*}(X) = T^{BR,*}(T^{BR}(T^*, X), X)$). Finally, define Home's and Foreign's utilities as functions of tariffs $T$ and $T^*$ and output $X$ by

$$w^{(*)}(T, T^*, X) \equiv u(c_1^{(*)}(p(T, T^*, X), T, X), c_2^{(*)}(p(T, T^*, X), T, X))$$

Notice that, since consumption (2), incomes (14), (15), and prices (16), are smooth functions of $T^{(*)}$ and $X$ and since $u$ is continuously differentiable the function $w$ is continuously differentiable as well.

With the notation thus established, the utilities under a trade agreement are $w^{(*)}(\bar{T}, \bar{T}^*, \bar{X})$, Home's and Foreign's defection utility on the trade agreement are, respectively, $w(T^{BR}(\bar{T}^*, \bar{X}), \bar{T}^*, \bar{X})$ and $w^{(*)}(\bar{T}, T^{BR,*}(\bar{T}, \bar{X}), \bar{X})$. Utilities in the first punishment phases are $w^{(*)}(T^N(\bar{X}), T^{N,*}(\bar{X}), \bar{X})$, while utilities $w^{N,(*)}$ of the second punishment phase – when tariffs and subsidies are set uncooperatively and react to the price changes – are independent of $\bar{T}$ and $\bar{X}$. Making use of this notation Home's self-enforcement constraint (21) of a time-invariant trade agreement can be written as

$$\begin{aligned} \frac{w(\bar{T}, \bar{T}^*, \bar{X})}{1-\beta} \geq\ & w(T^{BR}(\bar{T}^*, \bar{X}), \bar{T}^*, \bar{X})... \\ & ... + \frac{\beta - \beta^M}{1-\beta} w(T^N(\bar{X}), T^{N,*}(\bar{X}), \bar{X}) + \frac{\beta^M}{1-\beta} w^N \end{aligned} \tag{22}$$

while the self-enforcement constraint for Foreign is

$$\begin{aligned}\frac{w^*(\bar{T},\bar{T}^*,\bar{X})}{1-\beta} \geq\ & w^*(\bar{T},T^{BR,*}(\bar{T},\bar{X}),\bar{X})... \\ & ... + \frac{\beta-\beta^M}{1-\beta}w^*(T^N(\bar{X}),T^{N,*}(\bar{X}),\bar{X}) + \frac{\beta^M}{1-\beta}w^{N,*}\end{aligned} \tag{23}$$

For given output the best-response tariffs $T^{BR}(T^*,X)$ and $T^{BR,*}(T,X)$ generally involve lengthy expressions (see e.g., Kennan and Riezman (1988) or Devereux (1997)). It seems clear that incorporating subsidies and the reaction of decentralized firms in a game between governments can amount to a demanding task. But fortunately, Dixit (1987) establishes the existence of at least one Nash equilibrium in pure strategies, which is characterized by prohibitive tariffs on all sides. Applying this idea to the framework of the present model, the corresponding equilibrium consists of infinite tariffs $T = T^* = \infty$ and zero subsidies. The corresponding equilibrium is merely a replication of the respective autarkic economies. For the following analysis it is of little importance which of the (potentially many) equilibria prevails in the second – the long run – punishment phase and its choice is left open. The long-run punishment utilities will be simply referred to as $w^N$ and $w^{N,*}$ without further specification. It is, however, important to notice that none of the potential long-run equilibria is affected by tariffs and subsidies of the trade agreement and, in particular, by cooperation output $\bar{X}$.

## 3.2 Best-Response Tariffs

As increases in output require lengthy adjustment periods, the short- and medium-term policies can be calculated taken output structure as given. In particular, defection and medium term tariffs are computed under constant output. It is shown in the appendix that at fixed output $X = (x_1, x_2, x_1^*, x_2^*)$, the individually optimal tariffs that maximize the citizens' welfare are

$$T^{BR}(T^*,X) = \sqrt{\frac{\frac{x_1}{x_1^*}(T^*+1)/T^*+1}{\frac{x_2}{x_2^*}(T^*+1)+1}} \tag{24}$$

for Home and

$$T^{BR,*}(T,X) = \sqrt{\frac{\frac{x_2^*}{x_2}(T+1)/T+1}{(T+1)\frac{x_1^*}{x_1}+1}} \tag{25}$$

for Foreign. Expressions (24) and (25) have a number of noteworthy implications. First of all, they show that at constant $X$ the interior best-response tariffs are unique. Second, the unique tariffs $T^{BR}$ and $T^{BR,*}$ have singularities at $x_1^* = 0$ and $x_2 = 0$, respectively. More precisely, in the limit of diminishing import competing production, the exporter's best-response tariff grows unbounded. This is easily verified by the expressions above, which imply

$$\lim_{x_1^* \to 0} \sqrt{x_1^*} T^{BR} \in (0, \infty) \qquad \text{and} \qquad \lim_{x_2 \to 0} \sqrt{x_2} T^{BR,*} \in (0, \infty) \tag{26}$$

Third, (24) and (25) imply that, at constant output, the interior Nash Equilibrium of the tariff game is described by

$$T^N(X) = \sqrt{\frac{x_1 + x_1^*}{x_1^*} \frac{x_2^*}{x_2 + x_2^*}} \qquad \text{and} \qquad T^{N,*}(X) = \sqrt{\frac{x_2^* + x_2}{x_2} \frac{x_1}{x_1^* + x_1}} \tag{27}$$

These expression finally imply[9]

$$\lim_{x_1^* \to 0} \sqrt{x_1^*} T^N \in (0, \infty) \qquad \text{and} \qquad \lim_{x_2 \to 0} \sqrt{x_2} T^{N,*} \in (0, \infty) \tag{28}$$

and

$$\lim_{x_1^* \to 0} \sqrt{x_1^*}^3 \frac{dT^N}{dx_1^*} \in (-\infty, 0) \qquad \text{and} \qquad \lim_{x_2 \to 0} \sqrt{x_2}^3 \frac{dT^{N,*}}{dx_2} \in (-\infty, 0) \tag{29}$$

With the best response functions (24) and (25) the Nash tariffs (27) as well as the properties (26), (28), and (29) one can turn to the different components of the self-enforcement constraint and, in particular, the utilities that both countries derive under defection.

## 3.3 Defection Utilities

In the limit $x_1^* \to 0$ ($x_2 \to 0$) Home's (Foreign's) defection utility exhibits a singularity, the degree of which is classified by the following

**Claim 2** *If $T^*$ is constant, then Home's defection utility satisfies*

$$\lim_{x_1^* \to 0} \sqrt{x_1^*} \frac{d}{dx_1^*} w(T^{BR}(T^*, X), T^*, X) \quad \in \quad (-\infty, 0) \tag{30}$$

$$\lim_{x_2 \to 0} \frac{d}{dx_2} w(T^{BR}(T^*, X), T^*, X) \quad \in \quad \mathbb{R} \tag{31}$$

[9] See Kennan and Riezman (1988) of verify $T^N$ and $T^{N,*}$ by combining both expression with (24) and (25).

*Similarly, if $T$ is constant Foreign's defection utility satisfies*

$$\lim_{x_2 \to 0} \sqrt{x_2} \frac{d}{dx_2} w^*(T, T^{BR,*}(T,X), X) \quad \in \quad (-\infty, 0) \tag{32}$$

$$\lim_{x_1^* \to 0} \frac{d}{dx_1^*} w^*(T, T^{BR,*}(T,X), X) \quad \in \quad \mathbb{R} \tag{33}$$

**Proof.** See Appendix. ■

Conditions (30) and (32) of the claim show that the first unit of a country's import-competing output induces an unbounded loss of its trade partner's defection utility. Thus, small amounts of output of $X_1$ from Foreign are very effective in depressing Home's defection incentives. Intuitively, small amounts of $X_1$-production in Foreign heavily reduce Home's market power in its export market and strongly curb Home's one-shot gains from defection.

Compared to these strong effects, the impact of a country's first unit of import-competing output on its own defection utility are shown to be negligible (see (31) and (33)).

## 3.4 Punishment Utilities

As discussed above, punishment decomposes in two phases, reflecting the short and the long run after a defection. The strategies in the second phase (the long run) are set according to one of the (possibly many) equilibria of the tariff-plus-subsidy game. The according utilities, denoted by $w^{N,(*)}$ in (22) and (23), are independent of $\bar{X}$, which stands for the output structure under the trade agreement.

In the first punishment phase, however, the output structures of both economies are inherited form the trade agreement and hence the utilities shortly after a defection period do depend on $\bar{X}$. The properties of these punishment utilities will be analyzed next. Similarly to the defection utilities, these punishment utilities exhibit singularities at $x_1^* = 0$ and $x_2 = 0$. The degree of these singularities is classified by the following

**Claim 3** *Home's punishment utility in the first phase satisfies*

$$\lim_{x_2 \to 0} \sqrt{x_2} \frac{d}{dx_2} w(T^N(X), T^{N,*}(X), X) \quad = \quad \infty \tag{34}$$

$$\lim_{x_1^* \to 0} \sqrt{x_1^*} \frac{d}{dx_1^*} w(T^N(X), T^{N,*}(X), X) \quad \in \quad \mathbb{R} \tag{35}$$

*Foreign's punishment utility in the first phase satisfies*

$$\lim_{x_1^* \to 0} \sqrt{x_1^*} \frac{d}{dx_1^*} w^*(T^N(X), T^{N,*}(X), X) \quad = \quad \infty \tag{36}$$

$$\lim_{x_2 \to 0} \sqrt{x_2} \frac{d}{dx_2} w^*(T^N(X), T^{N,*}(X), X) \quad \in \quad \mathbb{R}. \tag{37}$$

**Proof.** See Appendix. ■

Equations (34) and (36) of the claim show that the adverse impact of the uncooperative tariffs in the first punishment phase can be substantially reduced by small amounts of import competing production. Indeed, comparing these equations with (30) and (32) show that the positive effect of the first unit of import competing production on a country's punishment utility is of higher degree than its adverse effect on the trade partner's defection utility. This qualitative difference is quite intuitive. Thus, at $x_1^* = 0$ Home can extract the maximum share of Foreign's income by driving its export price to infinity and its export quantity to zero. According to (30) small positive amounts of $x_1^*$ curb this ability, which, in terms of income change, hurts Home as much as it helps Foreign. In addition to this income effect, small amounts of the import competing $x_1^*$-production raise Foreign's consumption of $X_1$ from zero to strictly positive levels, which, by the standard properties of utility functions, adds infinite utility gains. This qualitative difference will be important for the design of the efficient trade agreement presented below.

Finally, (34) - (37) show that the first unit of import-competing production has a stronger impact on the country's medium-term punishment utility than on that of the trade partner.

Claims 2 and 3 describe the impact of small amounts of import-competing production on the utilities under defection and in the first punishment phase. With these assessments, one can turn to the main part of the paper, the optimal design of trade agreements.

## 3.5 Constrained Optimal Trade Agreements

The previous subsection has shown that the output structure strongly affects the defection temptation and hence the self-enforcement constraints. In particular, by generating import competing output under the marginal rates of transformation (18), subsidies may play a key role in optimal trade agreements. The following two subsections present the main results of the paper by characterizing the Pareto optimal time-invariant and self-enforceable trade agreements, which will, for brevity, be simply referred to as *optimal trade agreements*. Formally, a trade agreement will be defined by the tariffs and subsidies it implements, i.e., by the policy vector $(\bar{T}, \bar{T}^*, \bar{z}_2, \bar{z}_1^*)$.

Under these definitions, an optimal trade agreement $(\bar{T}, \bar{T}^*, \bar{z}_2, \bar{z}_1^*)$ maximizes the weighted sum the two countries' welfare subject to both self-enforcement constraints, i.e., $(\bar{T}, \bar{T}^*, \bar{z}_2, \bar{z}_1^*)$ solves the program

$$\max_{\bar{T}^{(*)}, \bar{z}_2, \bar{z}_1^*} \sigma w(\bar{T}, \bar{T}^*, \bar{X}) + (1-\sigma) w^*(\bar{T}, \bar{T}^*, \bar{X}) \quad s.t. \; (22) \text{ and } (23) \tag{38}$$

where $\sigma \in (0,1)$. The self-enforcement constraints of Home and Foreign will be rewritten as non-negativity restrictions on the functions $\Gamma^{(*)}$, which are defined, respectively, by

$$\begin{aligned} \Gamma = w(\bar{T}, \bar{T}^*, \bar{X}) - (1-\beta) w(T^{BR}(\bar{T}^*, \bar{X}), \bar{T}^*, \bar{X}) ... \\ ... - \left(\beta - \beta^M\right) w(T^N(\bar{X}), T^{N,*}(\bar{X}), \bar{X}) - \beta^M w^N \end{aligned} \tag{39}$$

and

$$\begin{aligned} \Gamma^* = w^*(\bar{T}, \bar{T}^*, \bar{X}) - (1-\beta) w^*(\bar{T}, T^{BR,*}(\bar{T}, \bar{X}), \bar{X}) ... \\ ... - \left(\beta - \beta^M\right) w^*(T^N(\bar{X}), T^{N,*}(\bar{X}), \bar{X}) - \beta^M w^{N,*} \end{aligned} \tag{40}$$

Combining these expression renders the Lagrangian of the optimization problem

$$\mathcal{L} = \sigma w(\bar{T}, \bar{T}^*, \bar{X}) + (1-\sigma) w^*(\bar{T}, \bar{T}^*, \bar{X}) + \lambda\Gamma + \lambda^*\Gamma^* + \nu \bar{z}_2 + \nu^* \bar{z}_1^* \tag{41}$$

where $\lambda \geq 0$ and $\lambda^* \geq 0$ stand for the Lagrange multipliers on Home's and Foreign's self-enforcement constraint and $\nu, \nu^* \geq 0$ are the Lagrange multipliers for the non-negativity constraints on $\bar{z}_2$ and $\bar{z}_1^*$, respectively.

It has been discussed at the start of this section that unconstrained efficient trade agreements implement a pair of tariffs satisfying $\bar{T}\bar{T}^* = 1$. Further, efficiency requires absence of wasteful subsidies ($\bar{z}_2 = \bar{z}_1^* = 0$), which, together with assumption (13), implies that full specialization emerges under the unconstrained optimal trade agreement.[10] As a point of reference, define $T_\sigma$ as Home's tariff of the unconstrained optimal trade agreement for $\sigma \in (0,1)$. Put differently, $T_\sigma$ solves the problem (38) in absence of (22) and (23). Define further $\bar{\beta}_\sigma$ as the minimum discount factor for which the self-enforcement constraints does not bind under $\bar{z}_i = \bar{z}_i^* = 0$, i.e.

$$\bar{\beta} \equiv \min\left\{\beta \in [0,1] \mid (22), (23) \text{ hold under } \left(\bar{T}, \bar{T}^*\right) = (T_\sigma, 1/T_\sigma) \text{ and } \bar{z}_i = \bar{z}_i^* = 0\right\}$$

[10] This can be verified with (9) and (16).

As the self-enforcement constraints is trivially satisfied in the limit $\beta \to 1$, $\bar{\beta} \in (0,1)$ holds. By construction of $\bar{\beta}$, the agreement $(\bar{T}, \bar{T}^*, \bar{z}_2, \bar{z}_1^*) = (T_\sigma, 1/T_\sigma, 0, 0)$ solves the constrained problem (38) for all $\beta \in [\bar{\beta}, 1]$. With these definitions one can formulate the first central result of the paper, which is summarized in the following proposition.

**Proposition 1** *There is an $\varepsilon > 0$ so that for all $\beta \in (\bar{\beta} - \varepsilon, \bar{\beta})$ the Pareto-optimal, time-invariant, self-enforcing trade agreement prohibits subsidies to import-competing sectors of the countries whose self-enforcement constraints bind. In these countries import-competing production is zero.*

**Proof.** Since $\beta < \bar{\beta}$ assume wlog $\lambda > 0$. Check with (9) and (16) that

$$\frac{a_2 L}{a_2^* L^*} = \frac{A_2 L}{A_2^* L^*} < TT^* < \frac{A_1 L}{A_1^* L^*} = \frac{a_1 L}{a_1^* L^*} \tag{42}$$

implies full specialization. Now define

$$\mathcal{T} = \{ (T, T^*) |\ T > T_\sigma,\ T^* > T_\sigma^*, \text{ and (42) holds} \}$$

as the set of tariff-pairs exceeding the unconstrained optimal tariffs consistent with full specialization. Equation (13) and $T_\sigma T_\sigma^* = 1$ imply $\mathcal{T} \neq \varnothing$. For $(T, T^*) \in \mathcal{T}$, equation (19) and the finite price $pT_\sigma$ under the trade agreement imply $dx_2/d\bar{z}_2 > 0$. Thus, the derivative of the Lagrangian (41) w.r.t. $\bar{z}_2$ is

$$\begin{aligned}
\frac{d\mathcal{L}}{d\bar{z}_2} \quad = \quad & \left[ (\sigma + \lambda) \tfrac{dw(\bar{T}, \bar{T}^*, \bar{X})}{dx_2} + (1 - \sigma + \lambda^*) \tfrac{dw^*(\bar{T}, \bar{T}^*, \bar{X})}{dx_2} \right] \frac{dx_2}{d\bar{z}_2} \dots \\
& \dots - (1 - \beta) \left[ \lambda \tfrac{dw(T^{BR}(\bar{T}^*, \bar{X}), \bar{T}^*, \bar{X})}{dx_2} + \lambda^* \tfrac{dw^*(\bar{T}, T^{BR,*}(\bar{T}, \bar{X}), \bar{X})}{dx_2} \right] \frac{dx_2}{d\bar{z}_2} \\
& \dots - \left( \beta - \beta^M \right) \left[ \lambda \tfrac{dw^*(T^N(\bar{X}), T^{N,*}(\bar{X}), \bar{X})}{dx_2} + \lambda^* \tfrac{dw^*(T^N(\bar{X}), T^{N,*}(\bar{X}), \bar{X})}{dx_2} \right] \frac{dx_2}{d\bar{z}_2} + \nu_2
\end{aligned}$$

Since $w$ is differentiable

$$\left| \frac{dw^{(*)}(\bar{T}, \bar{T}^*, \bar{X})}{dx_2} \right| < \infty$$

holds so that, by (31), (32), (34), (37) and $\lambda > 0$

$$\lim_{x_2 \to \infty} \frac{d\mathcal{L}}{d\bar{z}_2} - \nu_2 = -\infty \tag{43}$$

holds. This implies that there is a $\tilde{z} > 0$ so that $d\mathcal{L}/d\bar{z}_2 = 0$ implies either $\bar{z}_2 = 0$ or $\bar{z}_2 \geq \tilde{z}$ for all

$(T, T^*) \in \mathcal{T}$. By optimality of $\bar{z}_2 = 0$ the condition $\bar{z}_2 \geq \tilde{z}$ establishes a lower bound of the welfare loss $\Delta W' > 0$ under $\bar{z}_2 \geq \tilde{z}$ relative to the unconstrained optimal agreement $(T_\sigma, 1/T_\sigma, 0, 0)$.

Defining $W \equiv \sigma w(T, T^*, \bar{X}) + (1-\sigma) w^*(\bar{T}, \bar{T}^*, \bar{X})$ conditions

$$\left.\frac{dW}{dT}\right|_{(\bar{T},\bar{T}^*)=(T_\sigma,1/T_\sigma)} = 0 \qquad \text{and} \qquad \left.\frac{dW}{dT^*}\right|_{(\bar{T},\bar{T}^*)=(T_\sigma,1/T_\sigma)} = 0 \tag{44}$$

must hold by unconstrained optimality of $(T_\sigma, 1/T_\sigma)$. Under full specialization $c_1 = I/2 = x_1 T/(T+1)$ is increasing in $T$ (constant in $T^*$) and (with (16)) $c_2 = I/(2pT) = x_2^*/(T^*+1)$ is constant in $T$ (decreasing in $T^*$). Hence, with $X_0 = (x_1, 0, 0, x_2^*)$

$$\left.\frac{dw(T, T^*, X_0)}{dT}\right|_{(\bar{T},\bar{T}^*)=(T_\sigma,1/T_\sigma)} > 0 \qquad \text{and} \qquad \left.\frac{dw(T, T^*, X_0)}{dT^*}\right|_{(\bar{T},\bar{T}^*)=(T_\sigma,1/T_\sigma)} < 0$$

Therefore, one can define $\Delta, \Delta^* > 0$ so that

$$\left.\frac{dw(T_\sigma + \delta\Delta, 1/T_\sigma + \delta\Delta^*, X_0)}{d\delta}\right|_{\delta=0} = 0$$

and hence by (44)

$$\left.\frac{dw^*(T_\sigma + \delta\Delta, 1/T_\sigma + \delta\Delta^*, X_0)}{d\delta}\right|_{\delta=0} = 0$$

When defecting (with $T \to \infty$) Home's consumption is, according to (2) and (16) $c_1 = I/2 = x_1$ and $c_2 = I/(2pT) = 2x_2^*/(T^*+1)$ and hence its defection utility $w(T^{BR}(T^*, X_0), T^*, X_0)$ is strictly decreasing in $T^*$. Together, this implies

$$\left.\frac{d\Gamma}{d\delta}\right|_{\delta=0} = -(1-\beta)\frac{dw(T^{BR}(T^*, X_0), T^*, X_0)}{dT^*}\Delta > 0$$

for $\Gamma$ from (39). Similarly,

$$\left.\frac{d\Gamma^*}{d\delta}\right|_{\delta=0} = -(1-\beta)\frac{dw^*(T, T^{BR,*}(T, X_0), X_0)}{dT}\Delta^* > 0$$

holds. Thus, for $\delta \in (0, \bar{\delta})$ with $\bar{\delta} > 0$ small enough $\Gamma^{(*)}$ is strictly increasing in $\delta$ and, moreover, $(T_\sigma + \delta\Delta, 1/T_\sigma + \delta\Delta^*) \in \mathcal{T}$ holds. Since $\Gamma^{(*)}$ from (39) and (40) are continuous and strictly decreasing in $\beta$, this implies that for all $\delta \in (0, \bar{\delta})$ $\exists\, \varepsilon > 0$ so that $\Gamma^{(*)} \geq 0$ hold for $\beta \in (\bar{\beta} - \varepsilon, \bar{\beta})$ under the agreement $(T_\sigma + \delta\Delta, 1/T_\sigma + \delta\Delta^*, X_0)$.

Finally, let the welfare loss of the agreement $(T_\sigma + \delta\Delta, 1/T_\sigma + \delta\Delta^*, X_0)$ relative to unconstrained optimal agreement $(T_\sigma, 1/T_\sigma, 0, 0)$ be defined as $\Delta W(\delta)$. By incomes (14), (15), prices (16) consumption (2) and continuity of the utility $u$, $\Delta W(\delta)$ is increasing and continuous in $\delta$. Hence, there is a $\tilde{\delta} \in (0, \bar{\delta}]$ so that $\Delta W' > \Delta W(\tilde{\delta})$. With the corresponding $\tilde{\varepsilon} > 0$ this proves that $\bar{z}_2 = 0$ is optimal. Finally, $(T_\sigma + \delta\Delta, 1/T_\sigma + \delta\Delta^*) \in \mathcal{T}$ and $\bar{z}_2 = 0$ imply $x_2 = 0$. ■

Proposition 1 highlights the impact of subsidies on the efficiency of trade agreements. It shows that production subsidies to import-competing sectors are optimally zero if the respective self-enforcement constraint binds marginally. The intuition of this strong result is the following. Since subsidies to import-competing sectors affect output of some periods ahead, they generate domestic supply of the import good in the periods following potential defection. In so doing, they generate some degree of self-sufficiency, mitigate the consequences of defection for the subsidizing country, and thus reduce its dynamic costs of defection and increase its temptation to defect.[11] This is the dominant effect of production subsidies and the reason why they undermine the trade agreement. Subsidizing domestic production of the import good may thus be read as an offensive move to mitigate the consequences of a planned defection on trade agreements.

The proposition has shown that the effect described dominates other effects of subsidies that could, potentially, stabilize trade agreements. for an intuition that such effects actually exist, consider a situation where one country's – say Home's – self-enforcement constraint binds while Foreign's does not. In this case, subsidizing import-competing production in Foreign reduces Home's incentive to defect, as can be read from (30) At the same time, Foreign's subsidies do not necessarily bring its own self-enforcement to bind. Thus, subsidies in Home harm the trade agreement (Proposition 1 applies) while Foreign's subsidies relax Home's self-enforcement constraint and enhance the efficiency of the agreement. In this case, the optimal trade agreement does not implement general rules on subsidies but tailors individual subsidy schemes according to individual defection incentives.[12]

Proposition 1 has shown that production subsidies play an important role for optimal trade agreements whereas Bagwell and Staiger (2006) argue that efficient trade agreements shall not address them. Both results point in different directions but, nevertheless, do not contradict each other. In particular, Proposition 1 does not show that an optimal trade agreement requires explicit subsidy rules and hard-wire them in its legal text. To be precise, when the legal code of trade agreements needs to specify those and only those policies whose globally efficient level differs from the level,

[11] Sustainable cooperation under positive subsidies would necessarily entail higher import tariffs to reduce the defection incentives. These tariff increase would induce the efficiency loss relative to the optimal agreement.

[12] It may of course be questioned to what extend such tailored agreements may be implemented in practice.

which countries would choose individually. Clearly, if it is in the countries' individual interest to set zero subsidies, then subsidies need not be part of the trade agreement, in which case Proposition 1 is interesting *per se* but has no practical value for trade agreements.

Going back to the approach in Bagwell and Staiger (2006), trade agreements may specify simply the market access of the optimal trade agreement and, given this specification, the countries find it individually optimal not to subsidize their import competing sectors. In that case, a trade agreement that targets market access grants efficiency, while impeding subsidies does neither good nor harm. Such a situation would reflect the spirit of the main message in Ederington (2001). It will be shown next that this is not the case.

### 3.6 GATT versus WTO Legislation

Bagwell and Staiger (2006) recently analyze the differences between the GATT and the WTO subsidy rules. They find that the current WTO legislation, which essentially impedes production subsidies, generates inefficiencies, while the former GATT rules that addressed market access only, had been efficient. Proposition 1 of the previous subsection points in a different direction without explicitly contradicting this finding. To conduct a direct comparison between the GATT and WTO legislation in the present paper's setup, the respective subsidy rules will be mapped in a reduced form to the current framework by the following

**Definition**

(i) WTO rules set $(\bar{T}, \bar{T}^{*}, \bar{z}_2, \bar{z}_1^{*})$ where subsidies are zero: $\bar{z}_2 = \bar{z}_1^{*} = 0$.

(ii) GATT rules requires Home's (Foreign's) policies not to decrease net import prices $p$ $(1/p)$.

These definitions substantially simplify the respective rules but capture their essence. On the one hand, the WTO Agreement on Subsidies and Countervailing Measures allows member countries to challenge and enforce the removal of other members' production subsidies. Since any increase in import competing output due to subsidies adversely affects the trade partner's relative export prices (compare (16)) all subsidies will be challenged within the present framework. Consequently, WTO subsidy rules *de facto* implement $\bar{z}_2 = \bar{z}_1^{*} = 0$.[13]

On the other hand, the GATT subsidy rules rely on the concept of market access. According to Bagwell and Staiger (2006) GATT rules are violated whenever a "government has bound a tariff in

[13] This argument holds for negligible costs of non-violation claims only.

a GATT negotiation [...] and then subsequently alters its domestic policies in a way that diminishes the market access implied by that original tariff negotiation." At given domestic policies, Home's (Foreign's) access to Foreign's (Home's) market is $M = x_1 - c_1$ ($M^* = x_2^* - c_2^*$). Export market access is thus determined through (2), (3), (5), and (14) for Home ((2), (3), (5), and (15) for Foreign) and hence by its net-of tariff import price $p$ ($1/p$). Therefore, an upper bound on import prices is a sufficient criterion to target for the GATT-type trade agreements.[14]

With this specification of the WTO and the GATT rules, a clear separation is possible between the finding of Proposition 1 and those of Bagwell and Staiger (2006).

**Proposition 2** *Assume $\beta \in (\bar{\beta} - \varepsilon, \bar{\beta})$ with $\varepsilon > 0$ from Proposition 1. Then, within the class of Pareto-optimal, time-invariant, self-enforcing trade agreements, GATT-type agreements are inefficient.*

**Proof.** Assume the trade agreement of the GATT-type is efficient. Let its world price be $\bar{p}$ and its underlying policies be denoted by $(\bar{T}, \bar{T}^*, \bar{z}_2, \bar{z}_1^*)$. Since $\beta < \bar{\beta}$ assume wlog that $\lambda > 0$ ($\lambda$ from (41)), *i.e.*, Home's constraint binds. Then, Proposition 1 implies $x_2 = \bar{z}_2 = 0$. Further, $\lambda > 0$ implies $\Gamma = 0$ with $\Gamma$ from (39). Observe that Home's net import price $p$ is continuously decreasing in $\bar{T}$ and $x_2$ and thus in $\bar{z}_2$ for $\bar{z}_2$ around zero. Hence, for each small $\delta > 0$ there are $\Delta_T$ and $\Delta_z$ with $0 < \Delta_T, \Delta_z < \infty$ so that the policies $(\bar{T} - \delta\Delta_T, \bar{T}^*, \delta\Delta_z, \bar{z}_1^*)$ imply the equilibrium price $p = \bar{p}$. By construction, Home's strategies $(\bar{T} - \delta\Delta_T, \delta\Delta_z)$ do not violate the trade agreement. At the same time, properties (30) and (34) imply the value of defection is strictly increasing in $\delta$, *i.e.*,

$$\frac{d}{d\delta}\left[w(T^{BR}(\bar{T}^*, \bar{X}), \bar{T}^*, \bar{X}) + \tfrac{\beta - \beta^M}{1-\beta} w(T^N(\bar{X}), T^{N,*}(\bar{X}), \bar{X}) + \tfrac{\beta^M}{1-\beta} w^N\right]_{\delta=0} = +\infty \qquad (45)$$

Thus, at $\delta > 0$ small enough, defecting on $(\bar{T} - \delta\Delta_T, \delta\Delta_z)$ renders Home strictly higher utility than defecting on $(\bar{T}, 0)$. Finally, since Home's self-enforcement constraint binds ($\Gamma = 0$) this implies that defection on $(\bar{T} - \delta\Delta_T, \delta\Delta_z)$ renders higher utility than respecting the agreement. This contradicts self-enforceability of the agreement and proves the statement. ■

Proposition 2 draws a clear division between Proposition 1 and the main findings in Bagwell and Staiger (2006). While the latter study argues that efficient trade agreements essentially target market access only, Proposition 2 shows that this limitation necessarily creates inefficiencies.[15]

[14] Notice that a reduction of tariff change that does not decrease the trade partner's market access is not considered a defection. In presence of policy redundancy, as in Bagwell and Staiger (2006), this difference is irrelevant since tariffs can be set and all effective choices are compensated by the remaining policies. The current specification nevertheless remains in accordance with the GATT legislation (see GATT (1986) Article XXIII; for a discussion and an interpretation of the code concerning the legal basis for disputes see also http://www.wto.org/english/tratop_e/dispu_e/disp_settlement_cbt_e/c4s2p1_e.htm).

[15] To be precise, Bagwell and Staiger (2006) concede that the WTO subsidy rules can improve upon GATT subsidy

Moreover, Proposition 2 shows that it is important through which policy mix a country grants market access to its trade partner. As long as there are two or more policy instruments, a given degree of market access is compatible with wide a range policy combinations, among which countries pick their individually optimal choice. Intuitively, these different combinations of policies affect defection temptations and thus invoke different measures to check these temptations. Clearly, the higher defection temptations, the higher the efficiency costs of taming them[16]. Within the scope of admissible policies, the highest possible defection incentives have to be accounted for by the self-enforcing trade agreement – a requirement that necessarily limits the degree of cooperation. On the contrary, when trade agreements constrain the policies directly, the policy combination with the least defection temptation can be singled out and implemented.

Interestingly, both findings, those of Bagwell and Staiger (2006) as well as Proposition 2, build on the fact that agreements, which target market access leave countries a degree of freedom to choose the optimal policy mix that is compatible with the specified market access. However, while Bagwell and Staiger (2006) stress that allowing countries to implement the individually preferred policy enhances global welfare, the current study points out that this freedom can introduce severe defection incentives and limit the degree of international cooperation.[17]

The key element that generates these diametrically opposed results is the self-enforcement requirement. Absent in Bagwell and Staiger (2006), it drives the results of the current study. In the current setup, subsidies impact the countries' defection incentives and hence the self-enforcement constraints via their effect on international production patterns. Since, moreover, the self-enforcement constraints need to be accounted for by the designers of the trade agreement (and not by individual countries) the subsidies have to be hard-wired into the legal code of trade agreements. The findings of the present paper thus contradicts those in Bagwell and Staiger (2006) and, at the same time, makes a case for the current WTO subsidy rules.

rules when policy instruments are limited and subsidies are of little importance for the countries' welfare. The two propositions of this paper, however, can easily be extended to the case where subsidies have a substantial positive effect on domestic welfare, as long as these welfare gains are linear in subsidies $\bar{z}_2$ and $\bar{z}_1^*$ and hence cannot overturn the sign of the relevant derivatives (43) and (45).

[16] In this particular model, socially optimal subsidies are zero. For positive subsidies, defection incentives can be checked by increasing tariffs. This introduces distortions that make the agreements less efficient than those impeding subsidies.

[17] In this particular example, socially optimal subsidies are zero. For positive subsidies, defection incentives can be checked by increasing tariffs. This introduces distortions that make the agreements less efficient than those impeding subsidies.

### 3.7 Discussion of the Results

The main results of the paper, formulated in Proposition 1 and 2 describe subsidy rules of Pareto optimal, time-invariant and self-enforcing trade agreements. They have been derived under a number of assumptions that may appear restrictive and which deserve a word of justification. At the same time, the relevance of the findings and their place in the literature is briefly discussed. This discussion proceeds along the list of the model's assumptions.

**Full Specialization.** Proposition 1 and 2 rely on full international specialization (assumption (13)), a potentially severe restriction. As discussed in connection with Proposition1, key to the mechanism is that, under free trade, each country covers its consumption of at least one good entirely by imports, and attains a certain degree of self-sufficiency in that sector by subsidizing. This type of international specialization arises in most trade models under a variety of setups (see, *e.g.*, Dornbusch et al (1977), Dornbusch et al (1980), Acemoglu and Ventura (2002), and Romalis (2004)). In light of this wider interpretation the findings require significantly milder degree of international specialization.

**Marginally Binding Constraints.** The propositions require that the discount factor $\beta$ falls marginally short of the level that grants an undistorted world economy (i.e., $\beta \in (\bar{\beta} - \varepsilon, \bar{\beta})$). Thanks to successive rounds of trade negotiations over the past decades, average import tariffs have substantially dropped and were about 5% around the turn of the century (see e.g., Subramanian and Weil (2007)). While some countries set tariffs to zero, most charge positive but moderate import tariffs. From the point of view of self-enforcing trade agreements, these observations suggest that trade agreements are almost, but not quite, unconstrained or, to put it differently most self-enforcement bind but only marginally. This observation is reflected by the second key assumption. – Further, it is worth to stress that, while shown to dominate for marginally constrained trade agreements, the beneficial effects of deliberate mutual economic dependence between countries may well work in other scenarios. *E.g.*, economic history illustrates its relevance: according to standard interpretation of the European Coal and Steel Community, its foundation relied on a mutual dependence on trade that raised the dynamic costs of defective actions and forested the adherence to international cooperation. In this historic case, the pooling of steel, coal, and, to some extent, wheat was meant to create a mutual dependence between the six Western European member nations and aimed to make cooperation indispensable (see Gillingham (1991)).

**Preferences.** A third important assumption concerns the preferences, which are assumed to give rise to constant expenditure shares. This assumption generates strong effects of import-competing output on best-response tariffs and thus on utilities. The set of goods with the key property – positive expenditure shares at unbounded prices – seems quite exclusive. Yet again, the findings should not be prematurely discounted. A pronounced dependence on imports and hence strong vulnerabilities to import shortages can arise from vertical international specialization and is not too farfetched an assumption. Thus, a strong dependence on import goods can arise between vertically specialized countries under increasingly unbundled international production chain. In particular, national dependence and vulnerability comes along with imports and exports of highly specialized and relation-specific intermediate goods, giving support to this third assumption.[18] – Finally, and more generally, one has to keep in mind that the model's strong assumptions are meant to exemplify the general principle that subsidies may play an important role in trade agreements. While changes to this specific setup can change the way that optimal subsidy-rules should be designed, but they will not affect the basic principle this paper aims to highlight.

**Time to Build.** The paper's findings crucially depend on the substantial time-lag with which output reacts to price jumps. More specifically, industry-wide output cannot increase from zero to positive numbers at short notice. The empirical studies by Montgomery (1995) and Koeva (2000) have been mentioned in support of this assumption. It should moreover be stressed that this assumption does not concern small adjustment in firm output but reflects time-to-build of plants and the set-up of entire industries. The according time-lags are likely to be far larger than capacity-building on an intra-firm project basis and are thus assumed to exceed the time required for tariff-changes.

In sum, the key preconditions of this paper's results, formulated in Propositions 1 and 2, give a rough but after all a not too exotic description of today's world economy and its trading system. Consequently, the paper's basic message is encouraging for the WTO subsidy legislation and its *de facto* prohibition of production subsidies. It argues, in particular, that production subsidies for import-competing sectors can be an indispensable part of the legal code of optimal trade agreements.[19]

With its theoretical support for a prohibition of the WTO subsidy rules, the present paper strongly contradicts the results of Bagwell and Staiger (2006), who argue that impeding subsidies in trade agreements "may ultimately do more harm than good to the multilateral trading system." The diametrically opposed results of these earlier findings and the present paper's go back to the nature

[18] Yi (2003) estimates substantial levels and growth rates of cross-border vertical integration.

[19] This interpretation of the results obviously follows the standard assumption that trade agreements need to explicitly address the self-enforceable rules and policies and that cooperation is not automatically achieved as in the case of tacit collusion.

of the self-enforcement requirement. First, the self-enforcement requirement does not constrain the policy choice of individual countries but the policies a trade agreement can successfully implement. Since, moreover, the self-enforcement constraint is strongly affected by the patterns of international specialization, which, in turn, are driven by production subsidies, the latter subsidies must be addressed and accounted for by those supranational institutions that write the legal code of trade agreements.

Similar to Bagwell and Staiger (2006), the present paper addresses the role of subsidies in trade agreements. In so doing, however, it departs in some important points from the earlier work. First, the key insights of Bagwell and Staiger (2006) rely on a degree of policy redundancy,[20] which is absent from the present paper's the model. In the present setting, however, there is no role for additional redundant policies as long as effective subsidizing is forbidden. In this case, additional policy instruments could not alter the effective optimal policies and would remain void (compare also footnote 13). Second, the present paper does not model non-violation claims under the GATT or WTO rules in detail. This choice has been made due to the complexity of both models – the multi-stage game in the setup of Bagwell and Staiger (2006) and the repeated game in the present one. As argued at the beginning of subsection 3.6, however, the mapping of the respective subsidy rules to the present paper's setting is reasonable. Third, and finally, the present paper abstracts entirely from all potential positive effects of production subsidies on national welfare. In particular, its model is stripped of the governments' objectives other than those maximizing welfare and the economic and political motives for non-trivial subsidies, which drive the results in the general setup of Bagwell and Staiger (2006), are assumed to be absent in the current approach. I do not apologize for the bias, however. It should go without saying that the aim of the current study is not to disprove the validity of Bagwell and Staiger (2006) but, instead, to present a mechanism that may justify the inclusion of subsidies in trade agreements and, in particular, that rationalizes the tough standing of the WTO subsidy rules. A balanced evaluation of the WTO rules will take both arguments into account and weigh their relative importance.

In a broader perspective, the paper stresses the role of mutual dependence in self-enforcing trade agreements. Thus, whenever mutual dependence prevails in periods of potential punishment, it increases the threat of punishment for the deviant country and effectively increases room for cooperation. (See also Staiger (1994) and Devereux (1997) on this point.) On the base of this argument, it might generally be argued that policies such as export subsidies that have the potential

[20] Linear dependence of consuption tax, subsidies, and tariffs generate policy redundance.

to enhance mutual dependence should be embraced. In this context, however, the targeting must be recalled, according to which output subsidies (or output taxes) should be used to target output in case mutual dependence is to be created through international specialization.[21] This principle, at last, justifies the paper's focus on production subsidies.

# 4 Conclusion

This paper uses the framework of self-enforcing trade agreements to analyze the role of subsidies in trade agreements. A number of new results emerge. First, trade negotiators who aim to maximize national welfare must address subsidies in their agreements. Second, production subsidies are either large or zero. Consequently, when self-enforcement constraints bind only marginally, the optimal trade agreements ban production subsidies. This last scenario, which is argued to be a rough, but fair, description of today's world economy, stands in stark contrast to the findings of Bagwell and Staiger (2006), who argue that the WTO's ban on all subsidies is too much of a good thing and the "WTO subsidy rules may ultimately do more harm than good to the multilateral trading system." In this sense the present paper's findings contradict the results presented in Bagwell and Staiger (2006). At the same time, it makes a strong case for the WTO subsidy rules, providing a useful complement for a balanced evaluation of the prevailing WTO legislation.

[21] See Bhagwati and Srinivasan (1969).

# A Appendix

**Proof. of the Best Response Tariffs (24) and (25).** Write the indirect utility of Home as $v(\pi, I)$ with $\pi = pT$ the price of Home's import good. $e(\pi, u)$ is the according expenditure function. Take derivatives w.r.t. $\pi$ of the identity $v(\pi, e(\pi, u)) = u$ to get $v_\pi + v_I e_\pi = 0$. (Subscripts stand for partial derivatives.) Combining this equation with the optimality condition for $T$ leads to

$$\frac{d}{dT}v(\pi, I) = v_I \frac{dI}{dT} + v_\pi \frac{d\pi}{dT} = v_I \left( \frac{dI}{dT} - e_{\pi_i} \frac{d\pi}{dT} \right) = 0$$

Using Shephard's Lemma ($e_\pi(\pi, u) = c_2$), and $c_2 = I/2\pi$ leads to[22]

$$\frac{d}{dT}\ln(I) = \frac{1}{2}\frac{d}{dT}\ln(\pi)$$

With $p$ from (16), $\pi = pT$, and $I = (Tx_1 + \pi x_2)/\tau$ from (14) this gives

$$\frac{x_1 - \pi x_2}{Tx_1 - \pi x_2} = \frac{T+1}{2}\frac{d}{dT}\ln(\pi)$$

With the shorthand $\mu_1 = 1 + (x_1^*/x_1)T^*\tau/\tau^*$ and $\mu_2 = 1 + (x_2^*/x_2)(T+1)/(T(T^*+1))$ rewrite Home's relative price $\pi = pT$ as $\pi = \mu_1 x_1/(\mu_2 x_2)$ and the optimality condition as

$$\frac{\mu_2 - \mu_1}{T\mu_2 - \mu_1} = \frac{T+1}{2}\frac{d}{dT}\ln(\pi)$$

With the derivatives $d\mu_1/dT = (\mu_1 - 1)/(T+1)$ and $d\mu_2/dT = -(\mu_2 - 1)/(T(T+1))$ the optimality condition is $0 = 1 - 1/\mu_1 + \frac{1}{T}(1 - 1/\mu_2) - 2(\mu_2 - \mu_1)/(T\mu_2 - \mu_1)$ or

$$\begin{aligned} 0 &= \left[\mu_2(\mu_1 - 1) + \frac{1}{T}\mu_1(\mu_2 - 1)\right](T\mu_2 - \mu_1) - \mu_1\mu_2(\mu_2 - \mu_1) \\ &= \left\{\mu_2 + \frac{\mu_1}{T}\right\}[T(\mu_1 - 1)\mu_2 - (\mu_2 - 1)\mu_1] \end{aligned}$$

implying $T(\mu_1 - 1)\mu_2 = (\mu_2 - 1)\mu_1$. With the definitions of $\mu_i$ this can be written as

$$T^2\left[(T^* + 1)x_2 + x_2^*\right] - \frac{x_2^*}{x_1^*}\left(x_1\frac{T^*+1}{T^*} + x_1^*\right) = 0$$

and proves the (24). Expression (25) follows by symmetry. ■

[22] The optimality condition thus depends on income and prices only and the solution presented in Kennan and Riezman (1988) generalizes to the present scenario. The rest of the proof is provided here for completeness.

**Proof of Claim 2.** Define Home's domestic relative prices as $\pi(T^*, X) = pT^{BR}(T^*, X)$, where $p$ is from (16) under $T = T^{BR}$. Use (2) to compute the partial derivatives $c_{1,I} = 1/2$, $c_{2,I} = 1/2\pi$, $c_{1,\pi} = 0$, and $c_{2,\pi} = -I/2\pi^2$. This implies

$$\frac{dw(T^{BR}, T^*, X)}{dx_i^{(*)}} = u_1(c_1, c_2)\left\{\frac{dI}{dx_i^{(*)}} - \frac{I}{2}\frac{d\ln(\pi)}{dx_i^{(*)}}\right\} = \frac{u_2(c_1, c_2)}{\pi}\left\{\frac{dI}{dx_i^{(*)}} - \frac{I}{2}\frac{d\ln(\pi)}{dx_i^{(*)}}\right\} \tag{A1}$$

For (A1) the optimality condition $u_2 = \pi u_1$ was used. By the Envelope Theorem $T^{BR}$ can be treated as a constant. The superscripts $^{BR}$ will be omitted in the following.

(i) Show (30). Use (14) to derive $dI/dx_1^* = 2x_2/(T+1)\,(d\pi/dx_1^*)$. Hence (A1) renders

$$\frac{dw(T^{BR}, T^*, X)}{dx_1^*} = u_1(c_1, c_2)\left\{\frac{2x_2}{T+1}\pi - \frac{I}{2}\right\}\left(\frac{d}{dx_1^*}\ln(\pi)\right)$$

Now use (16) to get

$$\frac{d\ln(\pi)}{dx_1^*} = \frac{T^*(T+1)}{x_1(T^*+1) + x_1^* T^*(T+1)} + \frac{\bar{A}^*(T+1)/T}{x_2(T^*+1) + x_2^*(T+1)/T}$$

where $\bar{A}^*$ is Foreign's finite marginal rate of transformation between $X_1$ and $X_2$. Combining these expressions implies with (14), (16), and (26)

$$\lim_{x_1^* \to 0} \frac{1}{T^{BR}} \frac{dw(T^{BR}, T^*, X)}{dx_1^*} = \left(\lim_{x_1^* \to 0} u_1(c_1, c_2)\right)\{-x_1\}\left(\frac{T^*}{x_1(T^*+1)}\right)$$

The limits $\lim_{x_1^* \to 0} c_i > 0$ imply $\lim_{x_1^* \to 0} u_1(c_1, c_2) \in (0, \infty)$. This proves (30); (32) follows by symmetry.

(ii) Show (31). Check with (16)

$$\begin{aligned}
\lim_{x_2 \to 0} \pi &= T\frac{x_1(T^*+1) + x_1^* T^*(T+1)}{x_2^*(T+1)} \\
\lim_{x_2 \to 0} \frac{d\ln(\pi)}{dx_2} &= \frac{-\bar{A}(T^*+1)}{x_1(T^*+1) + x_1^* T^*(T+1)} - \frac{T(T^*+1)}{x_2^*(T+1)}
\end{aligned}$$

where $T = \sqrt{x_1(T^*+1)/(x_1^* T^*) + 1}$ according to (24). Note that both expressions are bounded for all $x_1^* \geq 0$. Thus, (14) implies

$$\lim_{x_2 \to 0} I = \frac{2T}{T+1}x_1 \qquad \lim_{x_2 \to 0} \frac{dI}{dx_2} = \frac{2}{T+1}(-\bar{A}T + \pi)$$

where $\bar{A}$ is Home's finite marginal rate of transformation between $X_1$ and $X_2$. Again, both expressions are bounded for all $x_1^* \geq 0$. Combining these expressions with (A1) leads to

$$\lim_{x_2 \to 0} \frac{dw(T^{BR}, T^*, X)}{dx_2} = \left( \lim_{x_2 \to 0} u_1(c_1, c_2) \right) \left\{ 2\frac{-\bar{A}T + \pi}{T+1} - \frac{T}{T+1} x_1 \frac{d\ln(\pi)}{dx_2} \right\}$$

By the reasoning above the expression in the slanted brackets is bounded for all $x_1^* \geq 0$. Finally, since $c_1 = I/2$ and $c_1 = I/2\pi$ are both positive at $x_2 \to 0$, $\lim_{x_1^* \to 0} u_1(c_1, c_2)$ is finite. This proves (31); (33) follows by symmetry. ■

**Proof of Claim 3.** Define Home's domestic relative price as $\pi(T^N, T^{N,*}, X) = pT^N$ with $p$ from (16) under $T^{(*)} = T^{N,(*)}$. Equation (A1) still applies. By the Envelope Theorem $T^N$ can be treated as a constant. The superscripts $^N$ will be omitted in the following.

(i) Show (34). If $x_1^* > 0$ use (16) and (27) to compute

$$\begin{aligned}
\lim_{x_2 \to 0} \pi &= \frac{x_1 + x_1^*(T+1)}{x_2^*(T+1)} \lim_{x_2 \to 0} T^* \\
\lim_{x_2 \to 0} \frac{d\ln(\pi)}{dx_2} &= \frac{-\bar{A}(T^*+1)}{x_1(T^*+1) + x_1^* T^*(T+1)} - \frac{T(T^*+1)}{x_2^*(T+1)} \ldots \\
&\ldots + \left\{ \frac{x_1 + x_1^*(T+1)}{x_1(T^*+1) + x_1^* T^*(T+1)} - \frac{Tx_2}{x_2^*(T+1)} \right\} \lim_{x_2 \to 0} \frac{dT^*}{dx_2}
\end{aligned}$$

where $\bar{A}^*$ is Foreign's finite marginal rate of transformation between $X_1$ and $X_2$. With (27) and (29) the second expression implies

$$\lim_{x_2 \to 0} x_2 \frac{d\ln(\pi)}{dx_2} \in (-\infty, 0) \tag{A2}$$

Further, (14) and (27) lead to

$$\begin{aligned}
\lim_{x_2 \to 0} I &= \frac{2Tx_1}{T+1} \\
\lim_{x_2 \to 0} \frac{dI}{dx_2} &= \frac{2(\pi - \bar{A}T)}{T+1} + \frac{2}{T+1} \lim_{x_2 \to 0} \left( \frac{x_2}{\pi} \frac{d\ln(\pi)}{dx_2} \right) = \frac{2(\pi - \bar{A}T)}{T+1}
\end{aligned}$$

Combining these expressions with (A1) leads to

$$\lim_{x_2 \to 0} \sqrt{x_2} \frac{dw(T^N, T^{N,*}, X)}{dx_2} = \left( \lim_{x_2 \to 0} \frac{u_2(c_1, c_2)}{\sqrt{x_2}\pi} \right) \left\{ -\frac{Tx_1}{T+1} \lim_{x_2 \to 0} x_2 \frac{d\ln(\pi)}{dx_2} \right\}$$

With (A2) the term in the slanted brackets is finite and positive; by (16) and (27) $\lim_{x_2 \to 0} \sqrt{x_2}\pi$ is finite. Since finally, $\lim_{x_2 \to 0} c_1 = \lim_{x_2 \to 0} I/2 > 0$ and $\lim_{x_2 \to 0} c_2 = \lim_{x_2 \to 0} I/2\pi = 0$ leads to

$\lim_{x_2 \to 0} u_2(c_1, c_2) = \infty$, which shows (34) for $x_1^* > 0$.

For the case $x_1^* = 0$ take (14), (16) and (27) to check

$$\lim_{x_1^* \to 0} \pi = \frac{x_1(T^* + 1)}{x_2(T^* + 1) + x_2^*} \qquad \text{and} \qquad \lim_{x_1^* \to 0} I = 2x_1$$

Hence,

$$\begin{aligned} \frac{dI}{dx_2} &= -2\bar{A} \\ \frac{d\ln(\pi)}{dx_2} &= \frac{-\bar{A}}{x_1} - \frac{T^* + 1}{x_2(T^* + 1) + x_2^*} + \left( \frac{1}{T^* + 1} - \frac{x_2}{x_2(T^* + 1) + x_2^*} \right) \frac{dT^{N,*}}{dx_2} \end{aligned}$$

With (27) this last expression shows that (A2) holds in the case $x_1^* = 0$ as well. Thus

$$\lim_{x_2 \to 0} \sqrt{x_2} \frac{dw(T^N, T^{N,*}, X)}{dx_2} = \left( \lim_{x_2 \to 0} \frac{u_2(c_1, c_2)}{\sqrt{x_2}\pi} \right) \left\{ -x_1 \lim_{x_2 \to 0} x_2 \frac{d\ln(\pi)}{dx_2} \right\}$$

By (16) and (28) $\lim_{x_2 \to 0} \sqrt{x_2}\pi$ is finite. Finally, observe that $\lim_{x_2 \to 0} c_1 = x_1/2 > 0$ while $\lim_{x_2 \to 0} c_2 = \lim_{x_2 \to 0} x_1/2\pi = 0$. Thus, $\lim_{x_2 \to 0} u_2(c_1, c_2) = \infty$ This completes the proof of (34); (36) follows by symmetry.

(ii) Show (35). If $x_2 > 0$ equations (14) and (16) imply

$$\lim_{x_1^* \to 0} \pi = \frac{x_1(T^* + 1)}{x_2(T^* + 1) + x_2^*} \qquad \text{and} \qquad \lim_{x_1^* \to 0} I = 2x_1$$

and

$$\begin{aligned} \lim_{x_1^* \to 0} \frac{d\ln(\pi)}{dx_1^*} &= \left( \lim_{x_1^* \to 0} T \right) \frac{T^*}{x_1(T^* + 1)} + \frac{A^*}{x_2(T^* + 1) + x_2^*} \ldots \\ &\quad \ldots + \left( \frac{1}{T^* + 1} - \frac{x_2}{x_2(T^* + 1) + x_2^*} \right) \frac{dT^*}{dx_1^*} \end{aligned}$$

Now, (27) implies $dT^*/dx_1^* < \infty$ for $x_2 > 0$ so that

$$\lim_{x_1^* \to 0} \sqrt{x_1^*} \frac{d\ln(\pi)}{dx_1^*} = \left( \lim_{x_1^* \to 0} \sqrt{x_1^*} T \right) \frac{T^*}{x_1(T^* + 1)} \in (0, \infty)$$

With

$$\lim_{x_1^* \to 0} \frac{dI}{dx_1^*} = 2 \lim_{x_1^* \to 0} \frac{x_2}{T} \frac{d\pi}{dx_1^*} \in (0, \infty)$$

and (A1) this leads to

$$\lim_{x_1^* \to 0} \sqrt{x_1^*} \frac{dw(T^N, T^{N,*}, X)}{dx_1^*} = \left( \lim_{x_1^* \to 0} u_1(c_1, c_2) \right) \left\{ \left( 2 \lim_{x_1^* \to 0} \frac{x_2 \pi}{T} - \frac{x_1}{2} \right) \frac{d \ln(\pi)}{dx_1^*} \right\}$$

The expression in the slanted brackets is finite. Finally, $\lim_{x_1^* \to 0} c_1 = x_1 > 0$ shows $\lim_{x_1^* \to 0} u_1(c_1, c_2) \geq 0$, which proves (35) for $x_2 > 0$.

For the case $x_2 = 0$ take (14), (16) and (27) to check

$$\lim_{x_2 \to 0} c_1 = \lim_{x_2 \to 0} I/2 = T/(T+1) \qquad \text{and} \qquad \lim_{x_2 \to 0} c_2 = I/(2\pi) = 0$$

Hence,

$$\lim_{x_1^* \to 0} \sqrt{x_1^*} \frac{dw(T^N, T^{N,*}, X)}{dx_1^*} = u_1 \left( \tfrac{T}{T+1}, 0 \right) \lim_{x_1^* \to 0} \left\{ \frac{\sqrt{x_1^*}}{(T+1)^2} \frac{dT}{dx_1^*} \right\}$$

By (27) and (29) the term in slanted brackets is finite. As $u_1(c_1, c_2)$ is bounded for positive $c_1$ this proves (35); (37) follows by symmetry ■

# References


Acemoglu, D., Ventura, J., 2002. The World Income Distribution. Quarterly Journal of Economics 117 (2), 659-694.

Anderson, K., 2004. The Challenge of Reducing Subsidies and Trade Barriers. CEPR Discussion Papers No. 4592.

Bagwell, K,. Staiger, R.W., 1999. An Economic Theory of GATT. American Economic Review 89 (1), 215-248.

Bagwell, K,. Staiger, R.W., 2000. GATT-Think. NBER WP No. 8005.

Bagwell, K,. Staiger, R.W., 2006. Will International Rules on Subsidies Disrupt the World Trading System? American Economic Review 96 (3), 877-895.

Baldwin, R.E., 2006. Failure of the WTO Ministerial Conference at Cancun: Reasons and Remedies. The World Economy 29 (6), 677-696.

Bhagwati, J., Srinivasan, T.N., 1969. Optimal Intervention to Achieve non Economic Objectives. Review of Economic Studies 36 (1), 27-38.

Bond, E.W., Park, J.H., 2002. Gradualism in Trade Agreements with Asymmetric Countries. Review of Economic Studies 69 (2), 379-406.

Devereux, M.B., 1997. .Growth, Specialization and Trade Liberalization. International Economic Review (38), 565-585.

Dixit, A., 1987. Strategic Aspects of Trade Policy, in: Truman F. (Ed.), Advances in Economic Theory. Fifth World Congress, New York, pp. 329-362.

Dornbusch, R., Fischer, S., Samuelson, P.A., 1977. Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods. American Economic Review 67, 823-839.

Dornbusch, R., Fischer, S., Samuelson, P.A., 1980. Heckscher-Ohlin Trade Theory with a Continuum of Goods. Quarterly Journal of Economics 95 (2), 203-224.

Ederington, J,. 2001. International Coordination of Trade and Domestic Policies. American Economic Review 91 (5), 1580-1593.

Furusawa, T., Lai, E. 1999. Adjustment Costs and Gradual Trade Liberalization. Journal of International Economics 49, 333-361.

GATT: The General Agreement on Tariffs and Trade, Text of the General Agreement. Geneva, July 1986.

Gillingham, J., 1991. Coal, Steel, and the Rebirth of Europe 1945-1955. The Germans and French from the Ruhr Conflict to the Economic Community, New York; Cambridge University Press. 1991.

Kennan, J., Riezman, R., 1988. Do Big Countries Win Tariff Wars? International Economic Review 29 (1), 81-85.

Koeva, P.Y., 2000. The Facts about Time-to-Build. IMF WP No. 00/138.

McLaren J., 1999. The Role of Multilateral Institutions in International Trade Cooperation. American Economic Review 89 (1), 190-214.

Mayer, W., 1981. Theoretical Considerations on Negotiated Tariff Adjustments. Oxford Economic Papers 33 (1), 135-153.

McLaren J., 1997. Size, Sunk Costs, and Judge Bowker's Objection to Free Trade. American Economic Review 87 (3), 400-420.

Montgomery, M.R., 1995. 'Time-to-build' Completion Patterns for Nonresidential Structures, 1961-1991. Economic Letters 48, 155-163.

Park, J.H., 2000. International Trade Agreements Between Countries of Asymmetric Size. Journal of International Economics 50, 473-495.

Romalis J. 1997. Factor Proportions and the Structure of Commodity Trade. American Economic Review 94 (1), 67-97.

Staiger, R.W., 1994. A Theory of Gradual Trade Liberalization, in Levinsohn, J., Deardorff, A.V., Stern, R.M., (Eds.) New Directions in Trade Theory, Ann Arbor, MI.

Subramanian, A., Wei, S.J. 2007. The WTO Promotes Trade, Strongly but Unevenly. Journal of International Economics 72 (3), 151-175.

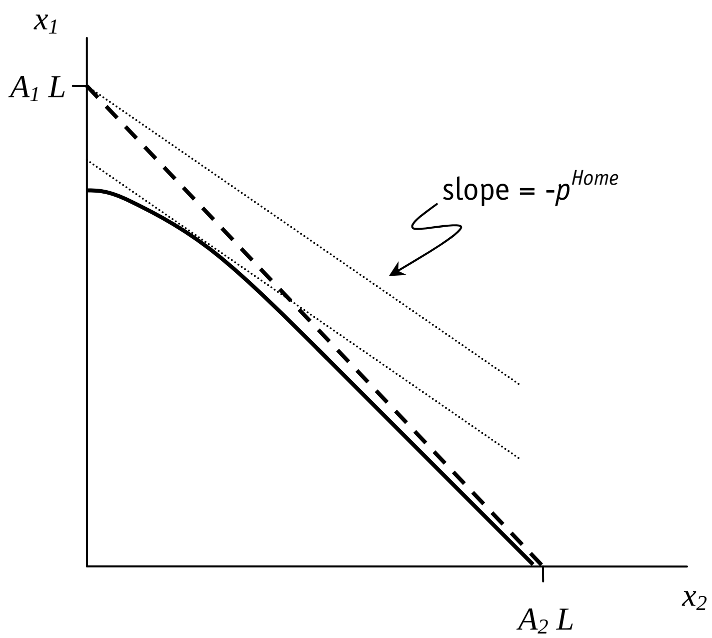
WTO 1995. Agreement Establishing the WTO, Annex 1, Agreement on Subsidies and Countervailing Measures. WTO, Geneva, Switzerland.

WTO 2007. Understanding the WTO. WTO, Geneva, Switzerland.

Yarbrough, B.V., Yarbrough,R.M., 1986. Reciprocity, Bilateralism, and Economic 'Hostages': Self-Enforcing Agreements in International Trade. International Studies Quarterly 30 (1), 7-21.

Yi, K.M., 2003. Can Vertical Specialization Explain the Growth of World Trade? Journal of Political Economy 111 (1), 52-102.

**Figure 1**




Production possibility frontier with subsidies (solid line) and without (dashed line).