"NASH-IN-NASH" TARIFF BARGAINING WITH AND WITHOUT MFN

Kyle Bagwell
Robert W. Staiger
Ali Yurukoglu

"Nash-in-Nash" Tariff Bargaining with and without MFN
Kyle Bagwell, Robert W. Staiger, and Ali Yurukoglu
NBER Working Paper No. 23894
October 2017
JEL No. F13

ABSTRACT

We provide an equilibrium analysis of the efficiency properties of bilateral tariff negotiations in a three-country, two-good general equilibrium model of international trade when transfers are not feasible. We consider "weak-rules" settings characterized by two cases: a no-rules case in which discriminatory tariffs are allowed, and an MFN-only case in which negotiated tariffs must be non-discriminatory (i.e., satisfy the MFN rule). We allow for a general family of political-economic country welfare functions and assess efficiency relative to these welfare functions. For the no-rules case with discriminatory tariffs, we consider simultaneous bilateral tariff negotiations and utilize the "Nash-in-Nash" solution concept of Horn and Wolinsky (1988). We establish a sense in which the resulting tariffs are inefficient and too low, so that excessive liberalization occurs from the perspective of the three countries. In the MFN-only case, we consider negotiations between two countries that are "principal suppliers" to each other and employ the Nash bargaining solution concept. Different possibilities arise. For one important situation, we establish a sense in which the resulting tariffs are inefficient and too high when evaluated relative to the unrestricted set of efficient tariffs. We also compare the negotiated tariffs under the MFN rule with the MFN-constrained efficiency frontier, finding that the negotiated tariffs are generically inefficient relative to this frontier and may lead to too little or too much liberalization. Finally, we illustrate our findings with a numerical analysis of a particular representation of the model as an endowment economy with Cobb-Douglas preferences and under the assumption that each government maximizes the indirect utility of the representative agent in its country.

Kyle Bagwell
Department of Economics
Stanford University
Landau Economics Building
579 Serra Mall
Stanford, CA 94305-6072
and NBER
kbagwell@stanford.edu

Robert W. Staiger
Department of Economics
Dartmouth College
319 Silsby
Hanover, NH 03755
and NBER
rstaiger@dartmouth.edu

Ali Yurukoglu
Graduate School of Business
Stanford University
Stanford, CA 94305
and NBER
ayurukog@stanford.edu

1 Introduction

Tariff negotiations between two countries can generate mutual gains by eliminating the inefficient terms-of-trade driven reductions in trade volume that occur under non-cooperative tariff setting. The effects of a bilateral trade deal, however, are not limited to the negotiating countries. Bilateral tariff cuts may also affect the welfare of other countries by altering their terms of trade. Due to this third-party externality, a tariff negotiation that is bilaterally efficient for the negotiating countries may fail to be efficient relative to the preferences of all countries.

By altering the terms of trade, bilateral tariff negotiations can affect both the export and import interests of third-party countries. To develop this point and explore the efficiency properties of bilateral tariff negotiations, we follow Bagwell and Staiger (2005, 2010, 2016) and consider a simple three-country, two-good model in which the home country exports a particular good to each of two foreign countries, where each foreign country in turn exports the other good to the home country and where the foreign countries do not trade with one another. In this setting, when the home country offers a tariff cut to one of the foreign countries, exporters in the other foreign country are disadvantaged and sell at a reduced world price. In addition, when a foreign country extends a tariff cut to the home country, the world price of the home country's export good increases, and so the other foreign country must pay a higher world price for its import good. A bilateral tariff liberalization thus induces a terms-of-trade loss for the third country, both by reducing world demand for that country's export good and by raising world demand for that country's import good.

Bagwell and Staiger (2005) show that, starting at any efficient vector of tariffs for the three countries, the home country and any one foreign country can always gain by extending bilateral tariff cuts to one another. Since the original tariffs are efficient, the bilateral tariff deal is necessarily opportunistic: the participating countries gain at the expense of the third-party foreign country, which suffers a terms-of-trade loss. This result suggests that the scope for bilaterally opportunistic trade deals is significant and indicates that an appropriately designed multilateral trade agreement can facilitate efficient outcomes for participating countries only if some restrictions are placed on the form of bilateral tariff deals. The GATT/WTO principle of non-discrimination, as captured by the most-favored nation (MFN) rule, can be motivated in this context. The MFN rule ensures the exporters from the non-participating foreign country enjoy any tariff cut that the home country offers as part of a bilateral deal. The MFN rule, however, does not fully insulate a given foreign country from the terms-of-trade effects of a bilateral negotiation between the home country and the other foreign country, both because in the presence of the MFN rule a tariff reduction by the home country now *raises* the world price of the good exported to the home country by both foreign countries, and because a tariff reduction by the other foreign country *still* raises the world price of the good imported from the home country by both foreign countries.[1]

---

[1]See Bagwell and Staiger (2005). They show, however, that the MFN rule when joined with the principle of reciprocity ensures that a bilateral tariff deal does not alter the terms of trade, nor thus the welfare, of the non-participating foreign country. For further discussion of the principle of reciprocity, see Bagwell and Staiger (1999,

Bagwell and Staiger (2005) develop their findings at a general level and do not study a specific extensive-form game of tariff bargaining among the three countries. They thus do not offer an equilibrium analysis of bargaining outcomes. In subsequent work, Bagwell and Staiger (2010) consider rules under which efficient outcomes can be achieved in a subgame perfect equilibrium of a sequential bargaining game for the three-country model when transfers are allowed, the MFN rule is required, and other restrictions on bilateral negotiations, including rules regarding reciprocity and renegotiation, may be imposed.[2] Bagwell and Staiger (2016) develop the analysis in a different direction, by characterizing the outcomes that can be achieved in a multilateral bargaining setting in which any proposed outcome must satisfy the MFN rule along with the principle of multilateral reciprocity. As they show, in this "strong-rule" setting, countries are unable to alter the terms of trade, and as a consequence multilateral bargaining outcomes may be characterized while requiring only that countries make dominant-strategy proposals.

In this paper, we provide an equilibrium analysis of the efficiency properties of bilateral tariff negotiations in the three-country model when transfers are not feasible and negotiations occur in "weak-rule" settings characterized by either no rules or only the MFN rule. We follow Bagwell and Staiger (2005, 2010, 2016) and assess efficiency relative to the preferences of countries, where these preferences are represented by general political-economic welfare functions that can include both economic and distributional concerns.[3] A basic feature of these weak-rule settings is that the negotiated tariffs in any one bilateral relationship may affect world prices and thus the payoffs (i.e., welfare levels) that are associated with negotiated tariffs in the other bilateral relationship. The bilateral negotiations are then fundamentally interdependent.

We begin with the case in which no rules are imposed. In the absence of even an MFN rule, the home country is free to negotiate bilateral agreements under which it applies discriminatory tariffs to exports from its two foreign trading partners. We may motivate consideration of the no-rules case in two ways. First, by comparing outcomes in the no-rules and MFN-only cases, we can gain insight into the implications of the MFN rule for tariff bargaining outcomes. Second, tariff discrimination is an important possibility in its own right, as it is a feature of various important historical trading relationships and arises to varying degrees in the current era as well among GATT/WTO member countries.[4]

2002, 2005, 2016) and Ossa (2014).

[2]See also Chan (2015).

[3]We thus include leading political-economy models of trade policy as well as the possibility that countries maximize national income. See Bagwell and Staiger (1999, 2002) for further discussion. For simplicity, in this paper, we refer to "government welfare" as "country welfare."

[4]See Pomfret (1997, Part 1) for a detailed history of discriminatory trade policies in the 19th and 20th centuries up to the early years of the WTO. As Beckett (1941, pp. 24-30), Pomfret (1997, p. 22), Rohlfing (2009), Tasca (1938, pp. 146-7) and Tavares (2006) discsus, tariff discrimination may occur even among trading relationships governed by the MFN rule, if countries limit the breadth of the MFN rule and impose narrow product reclassifications that enable the application of different tariffs on imports of broadly similar goods from different trading partners. Beckett (1941, p.28), for example, reports that approximately 1/3 of the tariff paragraphs modified by concessions negotiated by the United States under the Reciprocal Trade Agreement Act were associated with product reclassifications. Discrimination is also explicitly allowed under certain GATT/WTO rules. Under GATT Article XXIV, a subset of member countries may form a preferential trading agreement, provided that the participating countries go all the way to free trade on substantially all products that they trade with one another. As well, GATT/WTO rules allow

For the no-rules case, we characterize the efficiency properties of simultaneous bilateral tariff negotiations. Our approach is to adopt the solution concept of Horn and Wolinsky (1988). Originally developed to examine incentives for horizontal mergers in the presence of exclusive vertical relationships, the Horn-Wolinsky solution is now frequently used in the Industrial Organization literature to consider surplus division in bilateral oligopoly settings where externalities exist across firms and agreements.[5] The Horn-Wolinsky solution is sometimes referred to as a "Nash-in-Nash" solution, since it can be thought of as a Nash equilibrium between separate bilateral Nash bargaining problems. In the Horn-Wolinsky solution, any given bilateral negotiation results in the Nash bargaining solution taking as given the outcomes of the other negotiations.[6]

The primary benefit of the Nash-in-Nash approach is that it offers a simple means of characterizing simultaneous bilateral bargaining outcomes in settings with interdependent payoffs. Correspondingly, and as emphasized in the Industrial Organization literature, an important advantage of the Horn-Wolinsky solution is that it provides a tractable foundation for quantitative analyses in a wide range of applications where negotiations are interdependent. An important limitation of the Nash-in-Nash approach, however, is that it does not require that the solution be immune to multilateral deviations. The Nash-in-Nash approach is most directly interpreted in terms of a "delegated agent" model where a player (e.g., a firm in a merger analysis, or a country in a tariff negotiation) may be involved in multiple bilateral negotiations while relying on separate agents for each negotiation, where agents are unable to communicate with one another during the negotiation process. This interpretation may be strained in many settings of interest, including tariff negotiations, where within-negotiation communication between agents associated with the same player may be feasible.[7,8] The interpretation is arguably less strained, however, in settings with bargaining frictions such that opportunities for communication and coordination arise only after bilateral bargaining positions have hardened. On balance, we believe that the tractability advantages of the Horn-Wolinsky solution make it a potentially valuable tool, albeit only one such tool, for exploring

---

for preferences for developing countries and for plurilateral agreements in which a subset of countries extend certain benefits to one another.

[5]For example, Crawford and Yurukoglu (2012) and Crawford, Lee, Whinston, and Yurukoglu (2016) explore negotiations between cable television distributors and content creators, while Grennan (2013), Gowrisankaran, Nevo, and Town (2015), and Ho and Lee (2017) consider negotiations between hospitals, medical device manufacturers, and health insurers.

[6]The Nash-in-Nash approach is broadly related to the pairwise-proof requirements that are directly imposed in contracting equilibria (Cremer and Riordan, 1987) or indirectly implied under the requirement of "passive" beliefs in vertical contracting models (McAfee and Schwartz, 1994 and Hart and Tirole, 1990). See McAfee and Schwartz (1995) for further discussion.

[7]Collard-Wexler, Gowrisankaran and Lee (2016) develop micro-foundations for the Nash-in-Nash approach for negotiations that concern bilateral surplus division. The trade application that we consider here is different, however, in that negotiations are over tariffs (rather than lump-sum transfers) which have direct efficiency consequences.

[8]In their study of the GATT Torquay Round, Bagwell, Staiger and Yurukoglu (2017) highlight the impact of failed bargains between the United States and several British Commonwealth countries on other bilateral negotiations within the round. The Nash-in-Nash approach would not seem well-suited for a study of this behavior, for example. More generally, the Nash-in-Nash approach does not seem well-suited for a multilateral bargaining setting in which any proposed outcome must satisfy the MFN rule and the principle of multilateral reciprocity. As Bagwell and Staiger (2016) discuss, for a strong-rule setting wherein these requirements are strictly imposed, a home-country proposal for greater liberalization in one bilateral relationship is feasible only if the proposal calls for less liberalization in the other bilateral relationship.

the efficiency properties of bilateral tariff negotiations in various institutional environments. Our paper provides a theoretical foundation for such explorations.

In the context of the three-country tariff negotiation considered here, the Nash-in-Nash approach is captured with a representation in which the home country simultaneously negotiates with each foreign country, where the bargaining outcome in each bilateral negotiation is determined by the Nash bargaining solution and under the assumption that the Nash bargaining outcome will be successfully achieved in the other bilateral negotiation. If we were to interpret this approach in terms of a delegated agent model, then we might imagine that the home country sends one agent to negotiate with one foreign country and another agent to negotiate with the other foreign country, where the home-country agents each maximize a common home-country welfare function but are unable to communicate with each other during the course of their respective bilateral negotiations.

We begin our formal analysis by defining an interior Horn-Wolinsky solution when discriminatory tariffs are allowed (i.e., in the no-rules case). We then assume the existence of an interior Horn-Wolinsky solution and characterize its efficiency properties. For the setting in which discriminatory tariffs are allowed, we establish a sense in which the resulting tariffs are inefficient and too low, so that excessive liberalization occurs from the perspective of the three countries. Formally, we start at an interior Horn-Wolinsky solution and explicitly construct a perturbation under which all four tariffs (two tariffs for the home country, and one tariff for each foreign country) are raised in a manner that generates welfare gains for each of the three countries.

Having thus constructed a particular tariff-increasing perturbation that is sufficient for Pareto gains for all countries, we then consider the necessary features of any Pareto-improving tariff perturbation, where we again start with an interior Horn-Wolinsky solution for the no-rules case with discriminatory tariffs. Given that the model allows for four tariffs, and that each country has a direct interest in each of the four tariffs, we would not expect to find that Pareto gains are possible only if each individual tariff is perturbed toward a higher value. Indeed, a country experiences an externality from a bilateral negotiation to which it is not party if and only if its terms of trade are altered as a consequence of the combined effect of the tariff changes in that negotiation. Building on this perspective, we show that, if all countries enjoy weak welfare gains under a perturbation from an interior Horn-Wolinsky solution, then the perturbation cannot be characterized by "opportunistic" bilateral tariff changes in both bilateral relationships, where opportunistic bilateral tariff changes are bilateral tariff changes that impose a welfare (i.e., terms-of-trade) loss on the non-participating country. Using this finding, we further show that, if at least one country strictly gains under such a perturbation, then there must exist a bilateral relationship in which at least one tariff rises while the other tariff in that relationship can fall but only to a limited extent.

Based on these sufficient and necessary results, we conclude for the no-rules case with discriminatory tariffs that simultaneous bilateral tariff negotiations are associated with excessive liberalization when judged relative to the preferences of all countries. We are not aware of a previous equilibrium analysis that establishes this conclusion. It is, however, interesting to compare our results to those in a large literature that examines the possible third-party effects of preferential trading agree-

ments. This literature imposes a significant restriction on the family of discriminatory tariffs (so that trade is free among preferred partners) and then explores different questions such as whether such agreements facilitate or hinder the achievement of global free trade.[9] We include as a special case the possibility that countries maximize national income, and for this special case global free trade is of course efficient. Our finding of excess liberalization even for national-income-maximizing countries arises because we allow countries to pursue bilateral agreements in which they exchange discriminatory import subsidies and liberalize beyond free trade.[10]

Overall, our analysis of the no-rules case provides a set of efficiency characterizations for bilateral tariff negotiations while using the Horn-Wolinsky solution. The associated Nash-in-Nash bargaining model is a workhorse model in applied work in Industrial Organization that studies surplus division in bilateral oligopoly, and our work here provides a theoretical foundation for related applications in the context of bilateral tariff negotiations. Relative to work in Industrial Organization, a novel feature of our analysis is that we study bilateral relationships with two-way interactions (each country both sells to and buys from its trading partner). Our focus on efficiency is appropriate given our aim to study the welfare implications of bilateral tariff negotiations. We emphasize again, though, that the Horn-Wolinsky solution also has limitations, since the associated delegated agent interpretation is restrictive and may be more naturally applied in some settings than others.

We next consider the implications of the MFN rule for efficiency. In the MFN-only case, the home country is restricted to apply the same import tariff to exports from its two foreign trading partners. For this case, and in line with broad features of GATT/WTO practice, we assume that the home country negotiates with only one foreign country, referred to as the "principal supplier," and then extends any tariff change to the other foreign country on an MFN basis. For our two-good model, the resulting bargaining solution then amounts to a standard Nash bargaining solution between two countries, where the other foreign country is subject to the consequent home-country MFN tariff and leaves its own tariff unaltered.[11] For simplicity, we refer to the resulting solution as the MFN solution.

We first examine the efficiency properties of an interior MFN solution when efficiency is defined relative to the full space of (potentially discriminatory) tariffs. We show that different cases arise, depending on whether or not the home country would prefer more trade with its principal supplier country when taking as given the world price determined under the MFN solution. If the home

___

[9]See, for example, Saggi and Yildiz (2010) and Saggi, Yildiz and Woodland (2013) for analyses of the endogenous formation of preferential trading agreements.

[10]While import subsidies are not commonly observed, the Lerner symmetry theorem ensures that the effects of an import subsidy can be equivalently generated by an export subsidy. We also note that, in the context of free trade agreements that set tariffs to zero on substantially all trade, additional "deep integration" commitments in some cases might play a role similar to the role of import subsidies in our formal model.

[11]In our two-good model, when the home country negotiates only with its principal supplier, the delegated agent interpretation mentioned above is no longer relevent since by assumption only two countries actually negotiate. In Section 9, we also discuss a multiple-equilibrium issue that would arise in our two-good model were we to allow the home country to simultanously negotiate (over tariff bindings) with both foreign countries in the sense of the Horn-Wolinsky solution concept while also imposing the MFN rule. Finally, we note that the Horn-Wolinsky solution extends in a straightforward way to multiple-good settings when the MFN rule is imposed, if each country negotiates its tariff for any given import good only with a single principal supplier of that good.

country prefers greater trade in this sense, we can construct a perturbation from the MFN solution under which all countries gain and all four tariffs are reduced. For this case at least, we thus establish a sense in which the tariffs in the MFN solution are inefficient in the sense of being too high. More generally, our findings indicate that the MFN rule provides a partial counterbalance to the forces that result in inefficiently low tariffs when discriminatory tariffs are permitted.[12]

We next consider the efficiency properties of an interior MFN solution when efficiency is defined relative to the MFN-constrained efficiency frontier. Drawing on Bagwell and Staiger's (2005) characterization of the MFN-constrained efficiency frontier, we show that the MFN solution is generically inefficient relative to the MFN-constrained efficiency frontier and may lead to either too little liberalization or too much liberalization relative to MFN-constrained efficient levels.

Finally, we explore a particular representation of our model so as to concretely illustrate and further develop the themes described above. The representation that we consider is an endowment economy in which consumers have Cobb-Douglas preferences that weigh both goods equally. Under the assumption that each government maximizes the indirect utility of the representative agent in its country, we provide numerical characterizations of Nash tariffs, efficient tariffs, the interior Horn-Wolinsky solution, and the MFN solution. Among other findings, we verify that the interior Horn-Wolinsky solution exists for this representation.

The paper is organized as follows. Section 2 presents the basic three-country model of trade that we analyze. As we discuss there, we consider a general family of welfare functions for countries. Section 3 contains our definition of an interior Horn-Wolinsky solution for the setting with discriminatory tariffs. For the discriminatory setting, Section 4 contains our construction of a Pareto-improving perturbation relative to an interior Horn-Wolinsky solution. Section 5 provides related findings concerning the necessary features of Pareto-improving perturbations for this setting. Section 6 contains our constructions of Pareto-improving perturbations relative to an interior MFN solution, while Section 7 considers the relationship of the MFN solution to the MFN-constrained efficiency frontier. In Section 8, we provide numerical characterizations of our findings for an endowment economy with Cobb-Douglas preferences. Section 9 contains a brief discussion of extensions related to MFN bargaining beyond the principal supplier rule and also to the possibility of renegotiation. Section 10 concludes.

2 Trade Model

We employ the same three-country model of trade as studied by Bagwell and Staiger (2005).[13] In this section, we briefly summarize this model and highlight some of the features that we build upon in later sections.

The model features one home country and two foreign countries, which trade two goods, x

---

[12]As we discuss in Section 6, we are unable to provide necessary features of Pareto-improving perturbations for the MFN setting, since the MFN solution leaves unrestricted the tariff of the foreign country that is not the principal supplier.

[13]For further development of the model, see Bagwell and Staiger (1999, 2002, 2010, 2016).

and y, that are normal goods in consumption and produced in perfectly competitive markets under conditions of increasing opportunity costs. Each foreign country trades only with the home country, which imports good x from each of the two foreign countries in exchange for exports of good y. This trading structure implies that the home country is the only country that has the opportunity to set discriminatory tariffs. As usual, foreign country variables are denoted with an asterisk.

The home local relative price is denoted as $p \equiv p_x/p_y$, where p_x (p_y) is the local price in the home country of good x (y). The local relative price in foreign country $*i$, $i = 1, 2$, is likewise denoted as $p^{*i} \equiv p_x^{*i}/p_y^{*i}$. The ad valorem import tariff that the home country applies to exports of good x from foreign country $*i$ is denoted as t^i, and the ad valorem import tariff that foreign country $*i$ applies to exports of good y from the home country is denoted as t^{*i}. Throughout, we assume that tariffs are non-prohibitive. The world relative price for trade between the home country and foreign country $*i$ is denoted as $p^{wi} \equiv p_x^{*i}/p_y$. The world and local prices are related as $p = \tau^i p^{wi} \equiv p(\tau^i, p^{wi})$ and $p^{*i} = p^{wi}/\tau^{*i} \equiv p^{*i}(\tau^{*i}, p^{wi})$, where $\tau^i = 1 + t^i$ and $\tau^{*i} = 1 + t^{*i}$. The implied linkage relationship is then that $p^{wi} = [\tau^i/\tau^j]p^{wj}$. Under MFN tariffs, $\tau^1 = \tau^2$ and hence a single world price emerges: $p^{wi} \equiv p^w$ for $i = 1, 2$. By contrast, under discriminatory tariffs, $\tau^1 \neq \tau^2$ and so $p^{wi} \neq p^{wj}$.

In this model, foreign country $*i$'s terms of trade are given by p^{wi}. The home country's bilateral terms of trade with foreign country $*i$ are likewise defined as $1/p^{wi}$. The home country's multilateral terms of trade can then be defined using a trade-weighted average of its bilateral terms of trade. Formally, we define the home country's multilateral terms of trade as $1/T$, where

$$T(p^{*1}, p^{*2}, p^{w1}, p^{w2}) \equiv \sum_{i=1,2} s^{*i}(p^{*1}, p^{*2}, p^{w1}, p^{w2}) \cdot p^{wi}$$

with

$$s^{*i}(p^{*1}, p^{*2}, p^{w1}, p^{w2}) \equiv E^{*i}(p^{*i}, p^{wi})/ \sum_{j=1,2} E^{*j}(p^{*j}, p^{wj})$$

and where $E^{*i}(p^{*i}, p^{wi})$ denotes exports of good x from foreign country $*i$ to the home country. We assume that the share functions, $s^{*i}(p^{*1}, p^{*2}, p^{w1}, p^{w2})$, are continuously differentiable.

We observe that, under MFN tariffs, $T = p^{wi} \equiv p^w$; thus, the home country's bilateral and multilateral terms of trade are equal under MFN tariffs. Intuitively, under MFN tariffs, the home country's bilateral terms of trade are invariant across foreign trading partners, and so the home country's multilateral terms of trade takes this common value as well. A discriminatory tariff policy, on the other hand, implies that, for all i, $T \neq p^{wi}$. To see why, suppose that the home country imposes a higher tariff on exports from foreign country $*i$, so that $\tau^i > \tau^j$ for $j \neq i$. Since $p = \tau^i p^{wi}$ for $i = 1, 2$, it then follows that $p^{wi} < p^{wj}$, and so the home country enjoys a better bilateral terms of trade with foreign country $*i$. The home country's multilateral terms of trade is then $1/T$, where T is a trade-weighted average of the home country's bilateral terms of trade.

As discussed further in Bagwell and Staiger (2005), trade-balance and market-clearing condi-

tions may be stated using this notation.[14] Let the market-clearing world prices be denoted as $\widetilde{p}^{wi}(\boldsymbol{\tau})$, for $i = 1, 2$, where $\boldsymbol{\tau} \equiv (\tau^1, \tau^2, \tau^{*1}, \tau^{*2})$. We assume that $\widetilde{p}^{wi}$ is a continuously differentiable function. For the general case in which home may select discriminatory tariffs, we assume further that $\widetilde{p}^{wi}$ is increasing in τ^j, τ^{*j} and τ^{*i}, and is decreasing in τ^i. Thus, foreign country $*i$ suffers a terms of trade loss when τ^j and τ^{*j} drop via bilateral negotiations between the home country and foreign country $*j$, for $j \neq i$. For the case in which the home country selects MFN tariffs, with $\tau \equiv \tau^1 = \tau^2$, we write the market-clearing world price as $\widetilde{p}^w(\tau, \tau^{*1}, \tau^{*2}) \equiv \widetilde{p}^{wi}(\tau, \tau, \tau^{*1}, \tau^{*2})$. We assume for this case that $\widetilde{p}^w$ is increasing in τ^{*1} and τ^{*2} and is decreasing in τ, which is to say that home enjoys a terms of trade gain when it raises τ or when τ^{*i} drops for some $i = 1, 2$.

Our assumptions on market-clearing world prices $\widetilde{p}^{wi}$, $i = 1, 2$, are standard and easily motivated. Consider the setting with discriminatory tariffs. When a country raises the tariff that it applies to exports from another country, we expect that the resulting reduction in demand would make this good relatively less expensive on the world market; thus, it is natural to assume that $\widetilde{p}^{wi}$ is decreasing in τ^i and increasing in τ^{*i}. It is also natural to assume that $\widetilde{p}^{wi}$ is increasing in τ^j and τ^{*j}. Intuitively, foreign country $*i$ suffers a terms-of-trade loss when either of the tariffs τ^j and τ^{*j} in the other trade relationship are decreased, since it then encounters a resulting diminished world demand for its export good (when τ^j decreases) or an enhanced world demand for its import good (when τ^{*j} decreases). Finally, in the MFN setting, an increase in $\tau \equiv \tau^1 = \tau^2$ is expected to result in a decline in $\widetilde{p}^w$ if direct effects dominate indirect effects.

With the underlying model of trade defined in this way, we assume that the home-country and foreign-country $*i$ preferences, respectively, are continuously differentiable functions that can be written as $w(p, T)$ and $w^{*i}(p^{*i}, \widetilde{p}^{wi})$, with all prices evaluated at market-clearing levels and thus determined by the underlying vector of tariffs, $\boldsymbol{\tau}$.[15] The key assumption that we impose on preferences is that each country benefits from a terms of trade gain: $\partial w / \partial T < 0$ and $\partial w^{*i} / \partial \widetilde{p}^{wi} > 0$. As Bagwell and Staiger (1999, 2002) discuss, this assumption holds when countries maximize national income, and it is also satisfied by the leading political-economy models of trade policy.[16]

Notice that, for a given country, welfare is affected by changes in the tariffs of other countries only if those changes lead to a change in the given country's terms of trade, where for the home country the relevant terms-of-trade measure is its multilateral terms of trade. An externally generated terms-of-trade change affects a country's welfare directly and also indirectly, through the induced changes in local prices. For example, a change in τ^i affects w^{*j} only if it affects $\widetilde{p}^{wj}$, where

---

[14]As Bagwell and Staiger (2005) describe, production levels as well as the distribution and level of factor incomes in each country are determined by the local relative price in that country. Consumption levels in each country depend on the local relative price in the country as well as the country's (multilateral) terms of trade, where the latter along with the local price determine the country's tariff revenue, which is distributed to consumers. As anticipated above, each country's export and import functions then can be expressed as functions of that country's local relative price and (multilateral) terms of trade. From here, trade-balance and market-clearing conditions may be expressed in terms of relative local and world prices.

[15]With the market-clearing world prices determined as $\widetilde{p}^{wi}(\boldsymbol{\tau})$, the market-clearing local prices are given as $p(\tau^i, \widetilde{p}^{wi}(\boldsymbol{\tau}))$ and $p^{*i}(\tau^{*i}, \widetilde{p}^{wi}(\boldsymbol{\tau}))$ and are thus also determined by tariffs. Using the definition of T provided above, it is now straightforward to see that the market-clearing value for T can also be expressed as a function of tariffs.

[16]Political-economy models allow for distributional concerns. For simplicity, we refer to "government welfare" as "country welfare" in this paper.

any change in $\widetilde{p}^{wj}$ directly impacts w^{*j} and also indirectly impacts w^{*j} through the induced change in $p^{*j} = \widetilde{p}^{wj}/\tau^{*j}$. Following Bagwell and Staiger (2005), we assume above that the direct effect of a terms-of-trade change on a country's welfare can be signed: holding fixed a country's local price, a country's welfare rises when the country experiences an improvement in its terms of trade. We make no assumptions here, however, about the impact of a local-price change on a country's welfare.

Bagwell and Staiger (1999, 2002, 2005) define the *politically optimal tariffs* as the vector of tariffs that satisfies $\partial w/\partial p = 0 = \partial w^{*i}/\partial p^{*i}$ for $i = 1, 2$. With discriminatory tariffs allowed, this definition imposes three conditions on four tariffs, so that there are many politically optimal tariffs. The *politically optimal MFN tariffs* are the politically optimal tariffs for which $\tau^1 = \tau^2$. These tariffs are uniquely defined. Bagwell and Staiger (1999, 2002) show that politically optimal tariffs are efficient relative to the country welfare functions w, w^{*1} and w^{*2} if and only if the tariffs also satisfy the MFN rule.

We can now represent a country's welfare in reduced form as $W(\boldsymbol{\tau}) \equiv w(p, T)$ and $W^{*i}(\boldsymbol{\tau}) \equiv w^{*i}(p^{*i}, \widetilde{p}^{wi})$, where it follows under our assumptions that $W(\boldsymbol{\tau})$ and $W^{*i}(\boldsymbol{\tau})$ are continuously differentiable functions. In view of the local- and world-price effects of changes in tariffs, Bagwell and Staiger (2005) do not impose general restrictions on the relationships between tariffs and reduced-form country welfare functions. They do impose some additional structure, however, on these relationships when tariffs are efficient, where efficiency is evaluated relative to country welfare functions. Specifically, for efficient tariffs, Bagwell and Staiger (2005) assume that

$$
\begin{aligned}
\frac{\partial W}{\partial \tau^i} &> 0 \text{ and } \frac{\partial W^{*i}}{\partial \tau^{*i}} > 0 \\
\frac{\partial W}{\partial \tau^{*i}} &< 0 \text{ and } \frac{\partial W^{*i}}{\partial \tau^i} < 0 \\
\frac{\partial W^{*i}}{\partial \tau^{*j}} &> 0 \text{ and } \frac{\partial W^{*i}}{\partial \tau^j} > 0.
\end{aligned}
\tag{1}
$$

Notice that $\partial W^{*i}/\partial \tau^i < 0$ ensures that foreign country $*i$ suffers a welfare loss from an externally generated terms-of-trade loss. As Bagwell and Staiger (2005) also observe, it then follows that the inequalities in the third line of (1) are in fact implied; that is, $\partial W^{*i}/\partial \tau^i < 0$ implies $\partial W^{*i}/\partial \tau^{*j} > 0$ and $\partial W^{*i}/\partial \tau^j > 0$, since reductions in τ^i and increases in τ^{*j} or τ^j simply represent alternative external policy changes that lead to a terms-of-trade loss for foreign country $*i$.

Under the assumptions given in (1), Bagwell and Staiger (2005) show that, at any efficient tariff vector,

$$
-\frac{\partial W}{\partial \tau^{*i}} \Big/ \frac{\partial W}{\partial \tau^i} > -\frac{\partial W^{*i}}{\partial \tau^{*i}} \Big/ \frac{\partial W^{*i}}{\partial \tau^i} > 0 > -\frac{\partial W^{*j}}{\partial \tau^{*i}} \Big/ \frac{\partial W^{*j}}{\partial \tau^i}.
\tag{2}
$$

This means that, at any efficient tariff vector, the home country and foreign country $*i$ could lower τ^i and τ^{*i} in such a fashion as to enjoy mutual gains while imposing a terms-of-trade loss and indeed a welfare loss on foreign country $*j$. In effect, starting at any efficient tariff vector, the home country and foreign country $*i$ can move τ^i and τ^{*i} into a downward lens of mutual gain

while generating a welfare loss for foreign country $*j$. In this sense, when discriminatory tariffs are allowed, any efficient point is vulnerable to bilateral opportunism.

Figure 1 illustrates the efficient tariff vector in a graph with τ^i and τ^{*i} on the axes. As shown there, at an efficient tariff vector, the iso-welfare curves for the home country and foreign country $*i$ admit a downward lens of mutual gain. The gain that a tariff pair in the downward lens offers to these two countries, however, comes at the expense of foreign country $*j$, which suffers a terms-of-trade loss.[17] The existence of a downward lens suggests the possibility of excessive liberalization in a fully specified simultaneous bilateral bargaining game. This suggestion is incomplete, however, since a movement of one tariff pair into the downward lens for that pair would in turn shift or perhaps even eliminate the position of the downward lens for the other tariff pair. Bagwell and Staiger (2005) do not provide an equilibrium analysis of a simultaneous bilateral bargaining game and thus do not offer results concerning the efficiency properties of the resulting bargaining outcome. By comparison, a central objective of the current paper is to characterize the efficiency properties of the equilibrium outcomes of a fully specified model of simultaneous bilateral tariff bargaining.

The material presented in the section represents the modeling framework on which our analysis in subsequent sections builds. Each of the sections below, however, is self-contained as regards the additional structure that is placed on the manner in which tariffs affect reduced-form country welfare functions; in particular, in our analysis below, we do not maintain the assumption (1) and the associated characterization in (2) of efficient tariffs. Instead, we will impose additional structure on reduced-form country welfare functions explicitly and as needed in the analysis that follows. Bagwell and Staiger (2005) also consider MFN-efficient tariffs, which are the efficient tariffs under the restriction of the MFN rule: $\tau \equiv \tau^1 = \tau^2$. We postpone further discussion of this scenario and the relevant background findings until Section 7.

3 Horn-Wolinsky Solution

In this section, we define the Horn-Wolinsky solution for our trade application with simultaneous bilateral bargaining. We focus here on the case where the home country is free to use discriminatory tariffs.

To define the Horn-Wolinsky solution concept for our tariff bargaining application, we fix an initial tariff vector, $\boldsymbol{\tau}_0 \equiv (\tau_0^1, \tau_0^2, \tau_0^{*1}, \tau_0^{*2})$, which we take to be exogenous. One possibility is that this vector corresponds to the prior or "standing" agreements in each bilateral relationship. We also fix an exogenous bargaining power parameter, $\alpha \in (0, 1)$, which takes a larger value when the home country has greater bargaining power relative to the foreign countries. We are now in position to describe the endogenous determination of the tariff vector $\boldsymbol{\tau} \equiv (\tau^1, \tau^{*1}, \tau^2, \tau^{*2})$ through bilateral negotiations.

---

[17]As a general matter, when a country experiences a terms-of-trade loss as a consequence of a change in external policies, we cannot conclude that the country enjoys a welfare loss, since we must also consider the induced change in local prices. Under the assumptions given in (1), however, foreign country $*j$ experiences a reduction in its welfare when external policy changes result in a terms-of-trade loss.

Consider the bilateral negotiation between the home country and foreign country $*1$. Beginning from their initial tariffs τ_0^1 and τ_0^{*1} and taking τ^2 and τ^{*2} as given, the home country and foreign country $*1$ choose their Nash bargaining tariffs to solve

$$\max_{(\tau^1, \tau^{*1}) \in S} \Delta W^1(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}; \tau_0^1, \tau_0^{*1}) \cdot \Delta W^{*1}(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}; \tau_0^1, \tau_0^{*1}) \tag{3}$$

subject to

$$W(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}) \geq W(\tau_0^1, \tau_0^{*1}, \tau^2, \tau^{*2})$$
$$W^{*1}(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}) \geq W^{*1}(\tau_0^1, \tau_0^{*1}, \tau^2, \tau^{*2}),$$

where $S \equiv [\underline{\tau}, \overline{\tau}]^2$ with $(\underline{\tau}, \overline{\tau}) \in \Re^2$ and $0 < \underline{\tau} < \overline{\tau}$,

$$\Delta W^1(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}; \tau_0^1, \tau_0^{*1}) \equiv [W(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}) - W(\tau_0^1, \tau_0^{*1}, \tau^2, \tau^{*2})]^\alpha$$

and

$$\Delta W^{*1}(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}; \tau_0^1, \tau_0^{*1}) \equiv [W^{*1}(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}) - W^{*1}(\tau_0^1, \tau_0^{*1}, \tau^2, \tau^{*2})]^{1-\alpha}.$$

The bilateral negotiation between the home country and foreign country $*2$ is analogous. Beginning from their initial tariffs τ_0^2 and τ_0^{*2} and taking τ^1 and τ^{*1} as given, the home country and foreign country $*2$ choose their Nash bargaining tariffs to solve

$$\max_{(\tau^2, \tau^{*2}) \in S} \Delta W^2(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}; \tau_0^2, \tau_0^{*2}) \cdot \Delta W^{*2}(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}; \tau_0^2, \tau_0^{*2}), \tag{4}$$

subject to

$$W(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}) \geq W(\tau^1, \tau^{*1}, \tau_0^2, \tau_0^{*2})$$
$$W^{*2}(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}) \geq W^{*2}(\tau^1, \tau^{*1}, \tau_0^2, \tau_0^{*2}),$$

where

$$\Delta W^2(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}; \tau_0^2, \tau_0^{*2}) \equiv [W(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}) - W(\tau^1, \tau^{*1}, \tau_0^2, \tau_0^{*2})]^\alpha$$

and

$$\Delta W^{*2}(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}; \tau_0^2, \tau_0^{*2}) \equiv [W^{*2}(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}) - W^{*2}(\tau^1, \tau^{*1}, \tau_0^2, \tau_0^{*2})]^{1-\alpha}.$$

We may understand the inequality constraints in the respective programs as participation constraints. As captured by these constraints, for a given bilateral negotiation between the home country and foreign country $*i$, where $i = 1, 2$, if the negotiation results in disagreement, then the home country and foreign country $*i$ revert to the disagreement tariff pair (τ_0^i, τ_0^{*i}) for their bilateral relationship. Importantly, the home country and foreign country $*i$ negotiate under the assumption that the "other" bilateral negotiation (i.e., the bilateral negotiation between the home country and foreign country $*j$, where $j = 1, 2$ and $j \neq i$) delivers the tariff pair (τ^j, τ^{*j}), whether

the bilateral negotiation between the home country and foreign country $*i$ results in agreement or disagreement.

Given $S \equiv [\underline{\tau}, \overline{\tau}]^2$ with $(\underline{\tau}, \overline{\tau}) \in \Re^2$ and $0 < \underline{\tau} < \overline{\tau}$, and for $(\tau_0^1, \tau_0^{*1}, \tau_0^2, \tau_0^{*2}) \in S^2$ and $\alpha \in (0,1)$, we now say that a tariff vector $\tau_{hw} \equiv (\tau_{hw}^1, \tau_{hw}^{*1}, \tau_{hw}^2, \tau_{hw}^{*2}) \in S^2$ is a *Horn-Wolinsky solution* if $(\tau_{hw}^1, \tau_{hw}^{*1})$ solves (3) given $(\tau^2, \tau^{*2}) = (\tau_{hw}^2, \tau_{hw}^{*2})$ and if $(\tau_{hw}^2, \tau_{hw}^{*2})$ solves (4) given $(\tau^1, \tau^{*1}) = (\tau_{hw}^1, \tau_{hw}^{*1})$. In other words, $\tau_{hw} \equiv (\tau_{hw}^1, \tau_{hw}^{*1}, \tau_{hw}^2, \tau_{hw}^{*2}) \in S^2$ is a Horn-Wolinsky solution if it simultaneously solves the programs given in (3) and (4). The Horn-Wolinsky solution can thus be interpreted as a "Nash-in-Nash" solution, since each bilateral pair selects its Nash bargaining solution under the assumption that the other bargaining pair does as well.[18]

We next define an *interior Horn-Wolinsky solution* as a Horn-Wolinsky solution for which $\tau_{hw} \equiv (\tau_{hw}^1, \tau_{hw}^{*1}, \tau_{hw}^2, \tau_{hw}^{*2}) \in (\underline{\tau}, \overline{\tau})^4$ and the following optimization conditions are satisfied:

$$-\frac{\partial W}{\partial \tau^{*i}} \Big/ \frac{\partial W}{\partial \tau^i} = -\frac{\partial W^{*i}}{\partial \tau^{*i}} \Big/ \frac{\partial W^{*i}}{\partial \tau^i}, \text{ for } i = 1, 2, \tag{5}$$

where all derivatives are evaluated at τ_{hw}. The optimization conditions in (5) are implied by the first-order conditions for the optimization of programs (3) and (4) under the following sufficient conditions: $\tau_{hw} \in (\underline{\tau}, \overline{\tau})^4$; $\partial W/\partial \tau^{*i}, \partial W/\partial \tau^i, \partial W^{*i}/\partial \tau^{*i}$ and $\partial W^{*i}/\partial \tau^i$ are non-zero at τ_{hw}; and the participation constraints hold with slack at τ_{hw}.[19] As we discuss in more detail in the next section, an interior Horn-Wolinsky solution thus ensures that the tariff pair agreed upon in any bilateral negotiation is "bilaterally efficient."

4   Discriminatory Tariffs: Sufficient Conditions for Pareto Gains

In this section, we suppose that an interior Horn-Wolinsky solution exists for the model with discriminatory tariffs, and we establish a sense in which the resulting tariffs must be inefficient and too low. Specifically, we provide sufficient conditions under which it is possible to construct a particular perturbation where all countries gain by raising their tariffs.

To begin our analysis, we suppose that the no-rules model under simultaneous bilateral bargaining delivers an outcome, $\tau_{hw} \equiv (\tau_{hw}^1, \tau_{hw}^{*1}, \tau_{hw}^2, \tau_{hw}^{*2})$, where τ_{hw} is an interior Horn-Wolinsky solution. As above, we represent the welfare of each country as a function of the vector of tariffs. Given interiority, we know that each tariff pair, $(\tau_{hw}^i, \tau_{hw}^{*i})$, is bilaterally efficient, holding fixed the other

---

[18]In the Appendix, we further develop the Nash-in-Nash representation of the Horn-Wolinsky solution and show that the solution can be interpreted as a generalized Nash equilibrium for a generalized two-person game in which the objective of player i is to choose τ^i and τ^{*i} so as to maximize the Nash bargaining solution objective $\Delta W^i(\cdot) \cdot \Delta W^{*i}(\cdot)$ while satisfying the associated participation constraints for the bargaining negotiation between the home country and foreign country $*i$. The game and solution concept are "generalized," since, due to the participation constraints, player i's feasible strategy set is affected by the strategic choices of player j, for $i, j = 1, 2$ and $i \neq j$. Given this representation, we can then utilize Debreu's (1952, 1983) existence theorem and directly provide sufficient conditions for the existence of a Horn-Wolinsky solution. See also Dasgupta and Maskin (2015) for further discussion of Debreu's contribution.

[19]Formally, the participation constraints hold with slack at τ_{hw} if $W(\tau_{hw}) > \max\{W(\tau_0^1, \tau_0^{*1}, \tau_{hw}^2, \tau_{hw}^{*2}),$ $W(\tau_{hw}^1, \tau_{hw}^{*1}, \tau_0^2, \tau_0^{*2})\}$, $W^{*1}(\tau_{hw}) > W^{*1}(\tau_0^1, \tau_0^{*1}, \tau_{hw}^2, \tau_{hw}^{*2})$ and $W^{*2}(\tau_{hw}) > W^{*2}(\tau_{hw}^1, \tau_{hw}^{*1}, \tau_0^2, \tau_0^{*2})$.

tariff pair. In other words, we know that our solution resides on the bilateral efficiency loci:

$$-\frac{\partial W}{\partial \tau^{*i}} \Big/ \frac{\partial W}{\partial \tau^i} = -\frac{\partial W^{*i}}{\partial \tau^{*i}} \Big/ \frac{\partial W^{*i}}{\partial \tau^i}, \text{ for } i = 1, 2. \tag{6}$$

In analogy with the assumptions in Bagwell and Staiger (2005) for points on the efficiency frontier, we assume that, at the Horn-Wolinsky solution tariff vector τ_{hw}, the welfare impacts of tariff changes satisfy the following restrictions: for $i, j = 1, 2$ and $i \neq j$,

$$\begin{aligned}
\frac{\partial W}{\partial \tau^i} &> 0 \text{ and } \frac{\partial W^{*i}}{\partial \tau^{*i}} > 0 \\
\frac{\partial W}{\partial \tau^{*i}} &< 0 \text{ and } \frac{\partial W^{*i}}{\partial \tau^i} < 0 \\
\frac{\partial W^{*i}}{\partial \tau^{*j}} &> 0 \text{ and } \frac{\partial W^{*i}}{\partial \tau^j} > 0.
\end{aligned} \tag{7}$$

Under these assumptions, each country would like to increase its own tariff, each country does not want its export good to confront a higher tariff, and each foreign country $*i$ gains from an increase in either of the tariffs τ^j and τ^{*j} in the other bilateral trading relationship. In line with our discussion of (1) in Section 2, we further note that the assumption $\partial W^{*i}/\partial \tau^i < 0$ ensures that foreign country $*i$ experiences a welfare reduction from an externally generated terms-of-trade loss and therefore implies that $\partial W^{*i}/\partial \tau^{*j} > 0$ and $\partial W^{*i}/\partial \tau^j > 0$. In other words, the inequalities in the third line of (7) are in fact implied by the second inequality in the second line of (7).

Starting at any such Horn-Wolinsky solution as captured by (6), and under the assumptions given in (7), our claim now is that we can increase all four tariffs in a way that raises the welfare of all three countries. This directly suggests a local sense in which the Horn-Wolinsky tariffs are "too low" from an efficiency standpoint.

The idea of the perturbation builds from footnote 11 in Bagwell and Staiger (2005). Bagwell and Staiger (2005) consider an efficient tariff vector and suppose that the tangency condition in (6) holds between the home country and some foreign country $*i$. They then consider a two-step perturbation as illustrated in Figure 2. In the first step, they increase τ^i and τ^{*i} in a fashion that maintains W^{*i}. This corresponds to the movement from point A to point B in the figure. This first-step perturbation results in no change in W^{*i}, a first-order increase in W^{*j} and a second-order loss in W (due to the tangency between the iso-welfare curves of the home country and foreign country $*i$). The second step is then to increase τ^j and decrease τ^{*j} in a fashion that maintains W^{*i}. We illustrate this step in the figure with the movement from point C to point D. This second-step perturbation results in no change in W^{*i}, a first-order loss in W^{*j} and a first-order gain in W. If the second-step perturbation is small relative to the first-step perturbation, then the perturbation in total results in no change in W^{*i} and first-order gains in W^{*j} and W, which contradicts the original hypothesis of an efficient tariff vector.[20]

---

[20]Bagwell and Staiger (2005) use this argument to establish that an efficient tariff vector cannot be characterized by a tangency, such as illustrated in Figure 2. This argument is part of their proof that efficient tariff vectors must admit a downward lens, as depicted in Figure 1.

We want to consider here a similar perturbation, but there are three differences. First, we start with a situation in which the tangency condition (6) holds between the home country and both foreign countries. Second, we want to find a perturbation that generates welfare gains to each of the three countries. (By contrast, in the Bagwell-Staiger, 2005 perturbation just defined, W^{*i} is unchanged.) Third, we want to construct a perturbation under which all four tariffs are increased. (By contrast, in the single Bagwell-Staiger, 2005 perturbation just described, τ^{*j} is decreased.)

The key idea is to do two Bagwell-Staiger (2005) perturbations simultaneously, so that each foreign country plays the role of "foreign country $*j$" in one perturbation and thus emerges with a welfare gain in the combined perturbation. If for each perturbation the second-step adjustment is small in comparison to the first-step adjustment, then the combined perturbation will also call for a higher tariff from each foreign country. In other words, we will construct a combined perturbation such that, for each foreign country, the first-step tariff increase that it undertakes when playing the role of foreign country $*i$ exceeds the second-step tariff decrease that it undertakes when playing the role of foreign country $*j$.

We now develop a formal representation of this idea. Specifically, starting at a tariff vector that satisfies (6), and under the assumption (7), we consider the following perturbation:

$$d\tau^1 = d\tau^2 = \epsilon + \sigma \tag{8}$$

$$
\begin{aligned}
d\tau^{*1} &= (-\frac{\partial W^{*1}}{\partial \tau^1}/\frac{\partial W^{*1}}{\partial \tau^{*1}})\epsilon + (-\frac{\partial W^{*2}}{\partial \tau^1}/\frac{\partial W^{*2}}{\partial \tau^{*1}})\sigma \\
&= (-\frac{\partial W}{\partial \tau^1}/\frac{\partial W}{\partial \tau^{*1}})\epsilon + (-\frac{\partial W^{*2}}{\partial \tau^1}/\frac{\partial W^{*2}}{\partial \tau^{*1}})\sigma
\end{aligned}
\tag{9}
$$

$$
\begin{aligned}
d\tau^{*2} &= (-\frac{\partial W^{*2}}{\partial \tau^2}/\frac{\partial W^{*2}}{\partial \tau^{*2}})\epsilon + (-\frac{\partial W^{*1}}{\partial \tau^2}/\frac{\partial W^{*1}}{\partial \tau^{*2}})\sigma \\
&= (-\frac{\partial W}{\partial \tau^2}/\frac{\partial W}{\partial \tau^{*2}})\epsilon + (-\frac{\partial W^{*1}}{\partial \tau^2}/\frac{\partial W^{*1}}{\partial \tau^{*2}})\sigma
\end{aligned}
\tag{10}
$$

where the equalities in the second lines of (9) and (10) follow from the bilateral efficiency conditions (6) which the starting tariffs are assumed to satisfy, and where $\epsilon > 0$ and $\sigma > 0$ are both small. We give a further condition below concerning the relative magnitudes of ϵ and σ.

We can now compute the welfare differentials. For the home country, we get

$$
\begin{aligned}
dW &= (\frac{\partial W}{\partial \tau^1} + \frac{\partial W}{\partial \tau^2})(\epsilon + \sigma) + \frac{\partial W}{\partial \tau^{*1}}[(-\frac{\partial W}{\partial \tau^1}/\frac{\partial W}{\partial \tau^{*1}})\epsilon + (-\frac{\partial W^{*2}}{\partial \tau^1}/\frac{\partial W^{*2}}{\partial \tau^{*1}})\sigma] \\
&+ \frac{\partial W}{\partial \tau^{*2}}[(-\frac{\partial W}{\partial \tau^2}/\frac{\partial W}{\partial \tau^{*2}})\epsilon + (-\frac{\partial W^{*1}}{\partial \tau^2}/\frac{\partial W^{*1}}{\partial \tau^{*2}})\sigma].
\end{aligned}
$$

Thus,

$$dW = [\frac{\partial W}{\partial \tau^1} + \frac{\partial W}{\partial \tau^2} + \frac{\partial W}{\partial \tau^{*1}}(-\frac{\partial W^{*2}}{\partial \tau^1}/\frac{\partial W^{*2}}{\partial \tau^{*1}}) + \frac{\partial W}{\partial \tau^{*2}}(-\frac{\partial W^{*1}}{\partial \tau^2}/\frac{\partial W^{*1}}{\partial \tau^{*2}})]\sigma > 0,$$

14

where the inequality follows since $\sigma > 0$ and (7) holds at the original tariff vector.

For foreign country $*1$, we get

$$dW^{*1} = (\frac{\partial W^{*1}}{\partial \tau^1} + \frac{\partial W^{*1}}{\partial \tau^2})(\epsilon + \sigma) + \frac{\partial W^{*1}}{\partial \tau^{*1}}[(-\frac{\partial W^{*1}}{\partial \tau^1}/\frac{\partial W^{*1}}{\partial \tau^{*1}})\epsilon + (-\frac{\partial W^{*2}}{\partial \tau^1}/\frac{\partial W^{*2}}{\partial \tau^{*1}})\sigma]$$
$$+ \frac{\partial W^{*1}}{\partial \tau^{*2}}[(-\frac{\partial W^{*2}}{\partial \tau^2}/\frac{\partial W^{*2}}{\partial \tau^{*2}})\epsilon + (-\frac{\partial W^{*1}}{\partial \tau^2}/\frac{\partial W^{*1}}{\partial \tau^{*2}})\sigma]$$

Thus,

$$dW^{*1} = [\frac{\partial W^{*1}}{\partial \tau^2} + \frac{\partial W^{*1}}{\partial \tau^{*2}}(-\frac{\partial W^{*2}}{\partial \tau^2}/\frac{\partial W^{*2}}{\partial \tau^{*2}})]\epsilon + [\frac{\partial W^{*1}}{\partial \tau^1} + \frac{\partial W^{*1}}{\partial \tau^{*1}}(-\frac{\partial W^{*2}}{\partial \tau^1}/\frac{\partial W^{*2}}{\partial \tau^{*1}})]\sigma$$

As noted above, we can think of the perturbation here as a combination of two Bagwell-Staiger (2005) perturbations, which we might think of as Home-$*1$ and Home-$*2$ perturbations (with the designated foreign country playing the role of foreign country $*i$ in Bagwell and Staiger, 2005). The ϵ part of dW^{*1} is then the gain in W^{*1} from the Home-$*2$ step-1 increase in τ^2 and τ^{*2}, where there is no first-order effect on W^{*1} from the Home-$*1$ step-1 increase in τ^1 and τ^{*1}. The σ part of dW^{*1} is then the loss in W^{*1} from the Home-$*2$ step-2 increase in τ^1 and decrease in τ^{*1} to keep W^{*2} fixed, where by construction there is no effect on W^{*1} from the Home-$*1$ step-2 increase in τ^2 and decrease in τ^{*2} that keeps W^{*1} fixed.

Under our assumption that the initial tariff vector satisfies (7), the term in dW^{*1} that is multiplied by ϵ is positive while the term that is multiplied by σ is negative; therefore, if ϵ is large relative to σ in the specific sense that

$$\epsilon > \{\frac{-[\frac{\partial W^{*1}}{\partial \tau^1} + \frac{\partial W^{*1}}{\partial \tau^{*1}}(-\frac{\partial W^{*2}}{\partial \tau^1}/\frac{\partial W^{*2}}{\partial \tau^{*1}})]}{\frac{\partial W^{*1}}{\partial \tau^2} + \frac{\partial W^{*1}}{\partial \tau^{*2}}(-\frac{\partial W^{*2}}{\partial \tau^2}/\frac{\partial W^{*2}}{\partial \tau^{*2}})}\}\sigma,$$

then $dW^{*1} > 0$. An exactly symmetric argument holds for foreign country $*2$.

Allowing for $i = 1, 2$, we thus select $\epsilon > 0$ and $\sigma > 0$ such that

$$\epsilon > \max_{i,j=1,2,i\neq j}\{\frac{-[\frac{\partial W^{*i}}{\partial \tau^i} + \frac{\partial W^{*i}}{\partial \tau^{*i}}(-\frac{\partial W^{*j}}{\partial \tau^i}/\frac{\partial W^{*j}}{\partial \tau^{*i}})]}{\frac{\partial W^{*i}}{\partial \tau^j} + \frac{\partial W^{*i}}{\partial \tau^{*j}}(-\frac{\partial W^{*j}}{\partial \tau^j}/\frac{\partial W^{*j}}{\partial \tau^{*j}})}\}\sigma \tag{11}$$

Under (11), we may conclude that the perturbation raises the welfare of each country.

The remaining issue is to confirm that the perturbation increases each tariff. It is clear from (8) that $d\tau^1 = d\tau^2 > 0$. Referring to (9) and (10), we see for $d\tau^{*i}$ that the coefficient on ϵ is positive while that on σ is negative. Thus, we have that $d\tau^{*i} > 0$ for $i, j = 1, 2, i \neq j$ if and only if

$$\epsilon > \max_{i,j=1,2,i\neq j}[\frac{-\frac{\partial W^{*j}}{\partial \tau^i}/\frac{\partial W^{*j}}{\partial \tau^{*i}}}{\frac{\partial W^{*i}}{\partial \tau^i}/\frac{\partial W^{*i}}{\partial \tau^{*i}}}]\sigma. \tag{12}$$

We now summarize our finding in the following proposition:

Proposition 1 *Suppose the Horn-Wolinsky model under simultaneous bilateral bargaining delivers an interior Horn-Wolinsky solution, $\tau_{hw} \equiv (\tau_{hw}^1, \tau_{hw}^{*1}, \tau_{hw}^2, \tau_{hw}^{*2})$, captured by (6). Suppose at this tariff vector that (7) holds. For sufficiently small $\epsilon > 0$, $\sigma > 0$ satisfying (11) and (12), the perturbation defined in (8)-(10) raises all tariffs and generates welfare gains for all three countries.*

We note that welfare gains accrue to all countries without separately assuming that (12) holds. The role of (12) is simply to ensure that all tariffs are increased as part of the perturbation.[21]

We now have established that conditions exist under which, starting at any interior Horn-Wolinsky solution, all three countries can gain through a perturbation under which they all raise their tariffs. We thus have formalized an interpretation in which tariffs are inefficient in the sense of being too low, at any interior Horn-Wolinsky solution.

5 Discriminatory Tariffs: Necessary Conditions for Pareto Gains

Our results in the preceding section may be understood as providing sufficient conditions for Pareto gains through tariff increases; specifically, starting at an interior Horn-Wolinsky solution, we construct a particular perturbation under which all countries gain by raising their tariffs. In this section, we again start at an interior Horn-Wolinsky solution for the model with discriminatory tariffs, but we now examine the necessary conditions for perturbations that give Pareto gains. Our main finding is that, starting at an interior Horn-Wolinsky solution, if all countries enjoy weak welfare gains under a perturbation, then the perturbation cannot be characterized by "opportunistic" bilateral tariff changes in both bilateral relationships. We also show that, if at least one country strictly gains under such a perturbation, then there must exist a bilateral relationship in which at least one tariff rises while the other tariff in that relationship can fall but only to a limited extent. In this sense, our findings in this section reinforce the interpretation we formalize in the previous section that, at any interior Horn-Wolinsky solution of the model with discriminatory tariffs, tariffs are inefficiently low.

To formalize these arguments, we begin with some definitions. Let $\tau_{hw} \equiv (\tau_{hw}^1, \tau_{hw}^{*1}, \tau_{hw}^2, \tau_{hw}^{*2})$ denote an interior Horn-Wolinsky solution, where we assume again that (6) and (7) hold at this vector. Starting at τ_{hw}, we consider a perturbation $d\boldsymbol{\tau} \equiv (d\tau^1, d\tau^{*1}, d\tau^2, d\tau^{*2})$. It is convenient to decompose the perturbation into the bilateral tariff changes that are implied for each bilateral relationship, $(d\tau^1, d\tau^{*1})$ and $(d\tau^2, d\tau^{*2})$. For $i = 1, 2$, it is also convenient to define for the Home-$*i$ bilateral relationship a function $\widehat{\tau}^i$ that maps the tariff of foreign country $*i$ to the tariff that the home country applies on exports from foreign country $*i$. Our starting point is an interior Horn-Wolinsky solution, and so we assume that the function captures this solution: $\tau_{hw}^i = \widehat{\tau}^i(\tau_{hw}^{*i})$. To

---

[21]It can also be shown that (11) implies (12) under additional assumptions. For example, given our assumptions in (7), this implication holds in a symmetric setting, where a setting is symmetric if foreign countries $*1$ and $*2$ have symmetric welfare functions W^{*1} and W^{*2} and if tariffs are symmetric with $\tau^1 = \tau^2$ and $\tau^{*1} = \tau^{*2}$. To make this argument, we utilize that $\partial W^{*i}/\partial \tau^i + \partial W^{*i}/\partial \tau^j < 0$ in a symmetric setting under (7). Given (7), foreign country $*i$ suffers a reduction in welfare from an externally generated terms-of-trade loss, and it thus follows from our assumption that $\widetilde{p}^w(\tau, \tau^{*1}, \tau^{*2})$ is decreasing in τ that $\partial W^{*i}/\partial \tau^i + \partial W^{*i}/\partial \tau^j < 0$ in a symmetric setting. In Section 6, we explicitly utilize this implication in (17) in the context of our analysis of MFN tariffs.

ensure that the function $\widehat{\tau}^i$ also captures the perturbation as it relates to the Home-$*i$ bilateral relationship, we require further that $d\tau^i = [d\widehat{\tau}^i(\tau_{hw}^{*i})/d\tau^{*i}]d\tau^{*i}$. We can then represent the perturbation as changes in foreign tariffs, $d\tau^{*1}$ and $d\tau^{*2}$, with the corresponding changes in home tariffs captured as $d\tau^1 = [d\widehat{\tau}^1(\tau_{hw}^{*1})/d\tau^{*1}]d\tau^{*1}$ and $d\tau^2 = [d\widehat{\tau}^2(\tau_{hw}^{*2})/d\tau^{*2}]d\tau^{*2}$. Thus, for a given perturbation, the bilateral tariff changes in the Home-$*i$ bilateral relationship can be represented as $(d\tau^i, d\tau^{*i})$ where $d\tau^i = [d\widehat{\tau}^i(\tau_{hw}^{*i})/d\tau^{*i}]d\tau^{*i}$.

For $i, j = 1, 2$ with $i \neq j$, we now say that the perturbation entails an *opportunistic bilateral tariff change in the Home-$*j$ bilateral relationship* if the bilateral tariff change described by $(d\tau^j, d\tau^{*j})$ reduces the welfare of foreign country $*i$:

$$[\frac{\partial W^{*i}}{\partial \tau^{*j}} + \frac{\partial W^{*i}}{\partial \tau^j} \frac{d\widehat{\tau}^j(\tau_{hw}^{*j})}{d\tau^{*j}}]d\tau^{*j} < 0, \tag{13}$$

where $d\tau^{*j} \neq 0$ thus holds given (13). As a general matter, we note that an opportunistic bilateral tariff change in the Home-$*j$ bilateral relationship does not necessarily imply that the perturbation $d\boldsymbol{\tau}$ reduces the welfare of foreign country $*i$, since the perturbation includes as well the tariff changes $(d\tau^i, d\tau^{*i})$ in the Home-$*i$ bilateral relationship.

Let us now consider a perturbation that entails an opportunistic bilateral tariff change in both bilateral relationships. In other words, we consider now a perturbation for which (13) holds for $i, j = 1, 2$ and $i \neq j$. Each foreign country then suffers from the tariff changes that occur in the "other" bilateral relationship. We ask the following question: Starting at an interior Horn-Wolinsky solution, is it possible that such a perturbation can generate weak welfare gains for all countries? We argue next that the answer to this question is "no," from which it follows that a perturbation generating weak welfare gains for all countries necessarily has non-opportunistic bilateral tariff changes for at least one bilateral relationship.

To make this argument, let us suppose to the contrary that the perturbation satisfies (13) for $i, j = 1, 2$ with $i \neq j$ and yet generates weak welfare gains for all three countries. Consider now the welfare change under the perturbation for foreign country $*i$:

$$dW^{*i} = [\frac{\partial W^{*i}}{\partial \tau^{*j}} + \frac{\partial W^{*i}}{\partial \tau^j} \frac{d\widehat{\tau}^j(\tau_{hw}^{*j})}{d\tau^{*j}}]d\tau^{*j} + [\frac{\partial W^{*i}}{\partial \tau^{*i}} + \frac{\partial W^{*i}}{\partial \tau^i} \frac{d\widehat{\tau}^i(\tau_{hw}^{*i})}{d\tau^{*i}}]d\tau^{*i} \geq 0,$$

where the inequality follows from the assumption that the welfare change is non-negative for all countries. Under (13), we see that the first term in this expression is negative; thus, it follows that

$$[\frac{\partial W^{*i}}{\partial \tau^{*i}} + \frac{\partial W^{*i}}{\partial \tau^i} \frac{d\widehat{\tau}^i(\tau_{hw}^{*i})}{d\tau^{*i}}]d\tau^{*i} > 0.$$

Using $\partial W^{*i}/\partial \tau^i < 0$ under (7), we may rewrite this inequality equivalently as

$$[\frac{\partial W^{*i}}{\partial \tau^{*i}} / \frac{\partial W^{*i}}{\partial \tau^i} + \frac{d\widehat{\tau}^i(\tau_{hw}^{*i})}{d\tau^{*i}}]d\tau^{*i} < 0, \tag{14}$$

where the inequality in (14) holds for $i, j = 1, 2$ with $i \neq j$.

We consider next the welfare change under the perturbation for the home country. We find that

$$
\begin{aligned}
dW &= [\frac{\partial W}{\partial \tau^{*j}} + \frac{\partial W}{\partial \tau^j} \frac{d\widehat{\tau}^j(\tau_{hw}^{*j})}{d\tau^{*j}}]d\tau^{*j} + [\frac{\partial W}{\partial \tau^{*i}} + \frac{\partial W}{\partial \tau^i} \frac{d\widehat{\tau}^i(\tau_{hw}^{*i})}{d\tau^{*i}}]d\tau^{*i} \\
&= [\frac{\partial W}{\partial \tau^{*j}}/\frac{\partial W}{\partial \tau^j} + \frac{d\widehat{\tau}^j(\tau_{hw}^{*j})}{d\tau^{*j}}]d\tau^{*j}\frac{\partial W}{\partial \tau^j} + [\frac{\partial W}{\partial \tau^{*i}}/\frac{\partial W}{\partial \tau^i} + \frac{d\widehat{\tau}^i(\tau_{hw}^{*i})}{d\tau^{*i}}]d\tau^{*i}\frac{\partial W}{\partial \tau^i},
\end{aligned}
$$

where we use $\partial W/\partial \tau^i > 0$ for $i = 1, 2$ by (7). We now use the fact that an interior Horn-Wolinsky solution is bilaterally efficient and thus characterized by tangencies in each bilateral relationship. In particular, using (6), we now have that

$$
dW = \{[\frac{\partial W^{*j}}{\partial \tau^{*j}}/\frac{\partial W^{*j}}{\partial \tau^j} + \frac{d\widehat{\tau}^j(\tau_{hw}^{*j})}{d\tau^{*j}}]d\tau^{*j}\}\frac{\partial W}{\partial \tau^j} + \{[\frac{\partial W^{*i}}{\partial \tau^{*i}}/\frac{\partial W^{*i}}{\partial \tau^i} + \frac{d\widehat{\tau}^i(\tau_{hw}^{*i})}{d\tau^{*i}}]d\tau^{*i}\}\frac{\partial W}{\partial \tau^i} < 0,
$$

where the inequality follows since each term in curly brackets is negative by (14) and $\partial W/\partial \tau^i > 0$ for $i = 1, 2$ by (7). Finally, we note that $dW < 0$ is a contradiction to our assumption that the perturbation generates weak welfare gains for all countries.

The following proposition summarizes our finding:

Proposition 2 *Suppose the Horn-Wolinsky model under simultaneous bilateral bargaining delivers an interior Horn-Wolinsky solution, $\tau_{hw} \equiv (\tau_{hw}^1, \tau_{hw}^{*1}, \tau_{hw}^2, \tau_{hw}^{*2})$, captured by (6). Suppose at this tariff vector that (7) holds. Starting at this solution, a small perturbation $d\boldsymbol{\tau} \equiv (d\tau^1, d\tau^{*1}, d\tau^2, d\tau^{*2})$ generates weak welfare gains for all three countries only if the bilateral tariff change in at least one bilateral relationship is not opportunistic.*

Intuitively, if a perturbation from an interior Horn-Wolinsky solution entails opportunistic bilateral tariff changes for the Home-$*i$ bilateral relationship, then foreign country $*j$ can enjoy a weak gain under the perturbation only if it gains from the bilateral tariff changes in the Home-$*j$ bilateral relationship. For an interior Horn-Wolinsky solution, however, we know that the bilateral tariffs in the Home-$*j$ bilateral relationship are set in a bilaterally efficient manner; thus, foreign country $*j$ can gain from a change in the bilateral tariffs that it and the home country apply to each other only if the home country loses from this change. Continuing from here, if the home country is to enjoy a weak gain from the perturbation, then its loss in the Home-$*j$ bilateral relationship must be offset by a gain in the Home-$*i$ bilateral relationship. But by analogous reasoning, if the interior Horn-Wolinsky solution entails opportunistic bilateral tariff changes for the Home-$*j$ bilateral relationship, then foreign country $*i$ can enjoy a weak gain from the perturbation only if it, too, enjoys a gain in the Home-$*i$ bilateral relationship. Since the bilateral tariffs in the Home-$*i$ bilateral relationship are likewise set in a bilaterally efficient manner, however, it is not possible to find bilateral tariff changes for the Home-$*i$ bilateral relationship such that both the home country and foreign country $*i$ enjoy gains.

Proposition 2 identifies a key necessary feature for Pareto-improving perturbations. This result is of interest in its own right, but it also provides a stepping stone toward understanding the

necessary features of tariff changes that deliver Pareto gains. We thus conclude this section by exploring the implications of this proposition for the nature of the underlying tariff changes that a Pareto-improving perturbation must deliver.

To develop our findings, we first recall that foreign country $*i$ experiences a welfare change as a consequence of bilateral tariff changes in the Home-$*j$ bilateral relationship if and only if the bilateral tariff changes alter foreign country $*i$'s terms of trade, $\widetilde{p}^{wi}$. We now confirm that, under our existing assumptions, starting at an interior Horn-Wolinsky solution, foreign country $*i$ suffers a welfare loss when it faces an externally generated deterioration in its terms of trade; that is, we show that, at the interior Horn-Wolinsky solution tariff vector τ_{hw}, and for $i = 1, 2$,

$$\frac{d}{d\widetilde{p}^{wi}} w^{*i}(p^{*i}, \widetilde{p}^{wi}) > 0, \tag{15}$$

where we recall that $p^{*i} = (1/\tau^{*i})\widetilde{p}^{wi}$. To confirm that (15) necessarily holds under our existing assumptions, we observe from (7) that at the interior Horn-Wolinsky solution tariff vector τ_{hw}, and for $i = 1, 2$,

$$\frac{\partial W^{*i}}{\partial \tau^i} = [\frac{d}{d\widetilde{p}^{wi}} w^{*i}(p^{*i}, \widetilde{p}^{wi})][\frac{d\widetilde{p}^{wi}}{d\tau^i}] < 0.$$

Since $\widetilde{p}^{wi}$ is decreasing in τ^i, it is now evident that (15) holds.

Given (15) and for $i, j = 1, 2$ with $i \neq j$, we see that a perturbation entails an *opportunistic bilateral tariff change in the Home-$*j$ bilateral relationship* if and only if the bilateral tariff change described by $(d\tau^j, d\tau^{*j})$ generates a terms-of-trade loss for foreign country $*i$. Accordingly, by (15), the meaning of our proposition is that weak Pareto welfare gains are possible under a perturbation only if the bilateral tariff changes for at least one bilateral relationship result in a weak terms-of-trade gain for the foreign country that is not a member of that relationship.

We are now ready to explore the necessary features of the tariff changes that a Pareto-improving perturbation must deliver. Specifically, we consider a perturbation with two properties: it generates weak welfare gains for all countries, and at least one country actually gains under the perturbation. The latter property rules out the trivial possibility where all tariffs are unaltered. As just noted, the first property ensures that, for some $i, j = 1, 2$ with $i \neq j$, the bilateral tariff changes in the Home-$*j$ bilateral relationship generates a weak terms-of-trade gain for foreign country $*i$.

Consider such a perturbation. A first point is that the perturbation must entail a change in τ^j, τ^{*j} or both. To see why, suppose that the perturbation changes neither τ^j nor τ^{*j}. Given that, at an interior Horn-Wolinsky solution, τ^i and τ^{*i} are set in a bilaterally efficient manner for the Home-$*i$ bilateral relationship, any change in τ^i and τ^{*i} must result in a welfare loss for the home country or foreign country $*i$. Furthermore, if τ^i and τ^{*i} were also unaltered, then the perturbation would fail to generate an actual welfare gain for any country. We conclude that the assumed properties of the perturbation necessitate a change in τ^j and/or τ^{*j}. Since $\widetilde{p}^{wi}$ is increasing in τ^j and τ^{*j}, we may further observe that foreign country $*i$ cannot enjoy the assumed weak terms-of-trade gain if τ^j and τ^{*j} are both reduced. It follows that τ^j rises, τ^{*j} rises, or both τ^j and τ^{*j} rise.

We may thus conclude the section with the following proposition:

Proposition 3 *Suppose the Horn-Wolinsky model under simultaneous bilateral bargaining delivers an interior Horn-Wolinsky solution, $\tau_{hw} \equiv (\tau_{hw}^1, \tau_{hw}^{*1}, \tau_{hw}^2, \tau_{hw}^{*2})$, captured by (6). Suppose at this tariff vector that (7) holds. Starting at this solution, a small perturbation $d\boldsymbol{\tau} \equiv (d\tau^1, d\tau^{*1}, d\tau^2, d\tau^{*2})$ generates a welfare gain for at least one country while not lowering the welfare of any other country only if at least one tariff rises.*

The proposition establishes that the described Pareto improvement requires an increase in at least one tariff, but it is important to recognize that the underlying argument also places restrictions on the extent to which other tariffs can fall. In particular, given (15), we know that a weak Pareto improvement requires that, in at least one bilateral relationship, the associated bilateral tariff changes generate a weak terms-of-trade gain for the non-member foreign country. As we argue above, if we assume further that the perturbation generates an actual welfare gain for at least one country, then we can conclude that at least one tariff in this bilateral relationship actually rises. The other tariff in this bilateral relationship may rise as well or it could fall. But if it falls, it cannot fall to such an extent as to reverse the weak terms-of-trade gain that the non-member foreign country must enjoy.

6 MFN Tariffs: Sufficient Conditions for Pareto Gains

We now suppose that the home-country tariffs satisfy the MFN rule, so that $\tau \equiv \tau^1 = \tau^2$. We assume for this setting that the home country engages in a bilateral bargain with only its principal supplier, which we take to be foreign country $*1$ for simplicity. The tariff that foreign country $*2$ applies is thus left untouched and remains fixed at some exogenous level, $\overline{\tau}^{*2}$.[22] Our goal in this section is to define the bargaining solution for this MFN setting and then characterize the efficiency of the resulting MFN solution when efficiency is defined relative to the full space of tariff policies, $\boldsymbol{\tau} \equiv (\tau^1, \tau^{*1}, \tau^2, \tau^{*2})$.

We show that, in one important case, all countries can gain starting at the MFN solution when all four tariffs are reduced in an appropriate fashion. For this case, we thus formalize an interpretation in which tariffs are inefficient in the sense of being too high at the MFN solution. This finding contrasts interestingly with Proposition 3, where we show that, starting at the Horn-Wolinsky solution for the model with discriminatory tariffs, a perturbation can generate a welfare gain for at least one country without lowering the welfare for any other country only if at least one tariff rises. More generally, our findings in this section indicate that the MFN rule works as a partial counterbalance to the forces that arise when discriminatory tariffs are permitted and which lead tariffs to be inefficiently low at any interior Horn-Wolinsky solution.

---

[22]In Section 9, we discuss an alternative bargaining setting in which the home country simultaneously negotiates (over tariff bindings) with foreign countries $*1$ and $*2$ in the sense of the Horn-Wolinsky solution concept while also satisfying the MFN rule.

Formally, with the home-country tariff satisfying the MFN rule so that $\tau \equiv \tau^1 = \tau^2$, and with foreign country $*2$'s tariff fixed at an exogenous level, $\overline{\tau}^{*2}$, we assume that the negotiation between the home country and foreign country $*1$ is captured by a Nash bargaining solution, with bargaining parameter $\alpha \in (0, 1)$. We refer to the associated outcome as the *MFN solution,* and we represent this outcome with the tariff vector $\boldsymbol{\tau}_m \equiv (\tau_m, \tau_m^{*1}, \tau_m, \overline{\tau}^{*2})$, where we have used that $\tau_m \equiv \tau^1 = \tau^2$ at the MFN solution and that the tariff of foreign country $*2$ is fixed so that $\tau^{*2} \equiv \overline{\tau}^{*2}$. Assuming an interior solution, the resulting bargaining tariff pair, (τ_m, τ_m^{*1}), is bilaterally efficient, holding fixed τ^{*2} at $\overline{\tau}^{*2}$. In other words, the solution resides on the bilateral efficiency loci for the MFN setting:

$$-\frac{\partial W}{\partial \tau^{*1}} \Big/ \frac{dW}{d\tau} = -\frac{\partial W^{*1}}{\partial \tau^{*1}} \Big/ \frac{dW^{*1}}{d\tau}, \tag{16}$$

where the total derivatives reflect the fact that welfare functions are expressed as functions of the tariff vector $\boldsymbol{\tau} \equiv (\tau^1, \tau^{*1}, \tau^2, \tau^{*2})$ where in the MFN setting $\tau \equiv \tau^1 = \tau^2$.

We now assume that, at the MFN solution tariff satisfying (16) and given $\tau^{*2} \equiv \overline{\tau}^{*2}$, the welfare impacts of tariff changes for $i, j = 1, 2$ and $i \neq j$ satisfy (7). This assumption in turn implies that

$$\frac{dW^{*i}}{d\tau} < 0 \tag{17}$$

holds at the MFN solution tariff satisfying (16) and given $\tau^{*2} \equiv \overline{\tau}^{*2}$. The inequality in (17) indicates that foreign country $*i$ suffers a welfare loss when the home-country MFN tariff is increased, when we start at the MFN solution tariff satisfying (16) and fix $\tau^{*2} \equiv \overline{\tau}^{*2}$. The inequality in (17) is implied by (7) under our assumptions, since an increase in τ leads to a reduction in $\widetilde{p}^w$ and thus a terms-of-trade loss for foreign country $*i$, where under (7) foreign country $*i$ suffers a welfare reduction from an externally generated terms-of-trade loss.

Starting at the MFN solution as captured by (16), and under the assumption given in (7) with (17) then implied, we seek to identify conditions under which the MFN solution is inefficient relative to the full space of tariff policies, $\boldsymbol{\tau} \equiv (\tau^1, \tau^{*1}, \tau^2, \tau^{*2})$. Our analysis is organized around cases. As noted above, we identify one case under which, starting at the MFN solution, it is possible to reduce all tariffs in a fashion that generates Pareto gains.

Case 1   To establish our results, we distinguish between three cases. The first case captures the possibility that, at the MFN solution and holding fixed the induced MFN world price $\widetilde{p}^w(\tau_m, \tau_m^{*1}, \overline{\tau}^{*2})$, the home country would prefer more trade whereas foreign country $*1$ would prefer less trade. This is the case for which welfare gains accrue to all countries when all tariffs are reduced in an appropriate fashion.

We begin by expressing the conditions that define our first case:

$$-\frac{\partial W}{\partial \tau^{*1}} \Big/ \frac{dW}{d\tau} = -\frac{\partial W^{*1}}{\partial \tau^{*1}} \Big/ \frac{dW^{*1}}{d\tau} > -\frac{\partial W^{*2}}{\partial \tau^{*1}} \Big/ \frac{dW^{*2}}{d\tau}, \tag{18}$$

where the expressions are evaluated at the MFN solution and where total derivatives again reflect

the fact that $\tau \equiv \tau^1 = \tau^2$ in the MFN setting. The equality in (18) captures a tangency between the iso-welfare curves of the home country and foreign country $*1$ in a graph with τ on the y axis and τ^{*1} on the x axis, which we illustrate in the left panel in Figure 3. The inequality indicates that the iso-world-price curve, which gives combinations of τ and τ^{*1} that preserve $\widetilde{p}^w(\tau, \tau^{*1}, \tau^{*2})$ at the value obtained at the MFN solution, is flatter, as Figure 3 depicts. This ensures that, at the prevailing world price, the home country desires more trade whereas foreign country $*1$ prefers less trade. Finally, since foreign country $*2$ is impacted by changes in τ and τ^{*1} only insofar as those changes impact the world price, we may equivalently express the slope of the iso-world-price curve in terms of the slope of the iso-welfare curve for foreign country $*2$.

The idea of the perturbation is as follows. In step 1 of the perturbation, the home country and foreign country $*1$ lower τ and τ^{*1} in a fashion that maintains W^{*1}. We illustrate this step in Figure 3 with the movement from point A to point B. Under the first case, this change necessitates a terms-of-trade improvement for foreign country $*1$ (i.e., an increase in $\widetilde{p}^w$) as compensation for the increased trade volume. Due to the tangency condition (16), the home country then experiences only a second-order loss from this first-step adjustment. By contrast, foreign country $*2$ has the same terms-of-trade as foreign country $*1$ and thus enjoys a terms-of-trade gain, and thus a first-order welfare gain, from the first-step adjustment. In the second step of the perturbation, we utilize the full tariff space and raise τ^2 while lowering τ^{*2}, again so as to maintain W^{*1}. This step is illustrated in Figure 3 with the movement from point C to point D. This second change generates a first-order gain for the home country and a first-order loss for foreign country $*2$. If the second-step adjustment is sufficiently small in magnitude, however, then the home country and foreign country $*2$ enjoy first-order gains overall and all tariffs drop.[23] At this point, we could allow a very small increase in the tariff of foreign country $*1$ to ensure that all countries gain from a perturbation in which all tariffs are reduced.

We now formalize this idea. Starting at a tariff vector that satisfies (16) and (18), and under the assumption given in (7) with (17) then implied, we now consider the following perturbation:

$$d\tau^1 = -\epsilon \tag{19}$$

$$d\tau^2 = -\epsilon + \sigma \tag{20}$$

$$d\tau^{*1} = \left(-\frac{dW}{d\tau}\Big/\frac{\partial W}{\partial \tau^{*1}}\right)(-\epsilon) = \left(-\frac{dW^{*1}}{d\tau}\Big/\frac{\partial W^{*1}}{\partial \tau^{*1}}\right)(-\epsilon) \tag{21}$$

$$d\tau^{*2} = \left(-\frac{\partial W^{*1}}{\partial \tau^2}\Big/\frac{\partial W^{*1}}{\partial \tau^{*2}}\right)\sigma \tag{22}$$

where the equality in (21) follows from the bilateral efficiency condition (16) which the starting tariffs are assumed to satisfy, and where $\epsilon > 0$ and $\sigma > 0$ are both small. We give a further condition below concerning the relative magnitudes of ϵ and σ.

---

[23]Note that the rise in τ^2 in the movement from point C to point D is then small in comparison to the fall in τ^2 that is embedded in the fall of τ in the movement from point A to point B.

We can now compute the welfare differentials. For the home country, we get

$$dW = \frac{dW}{d\tau}(-\epsilon) + \frac{\partial W}{\partial \tau^2}\sigma + \frac{\partial W}{\partial \tau^{*1}}(-\frac{dW}{d\tau}/\frac{\partial W}{\partial \tau^{*1}})(-\epsilon) + \frac{\partial W}{\partial \tau^{*2}}(-\frac{\partial W^{*1}}{\partial \tau^2}/\frac{\partial W^{*1}}{\partial \tau^{*2}})\sigma.$$

Thus,

$$dW = [\frac{\partial W}{\partial \tau^2} + \frac{\partial W}{\partial \tau^{*2}}(-\frac{\partial W^{*1}}{\partial \tau^2}/\frac{\partial W^{*1}}{\partial \tau^{*2}})]\sigma > 0,$$

where the inequality follows since $\sigma > 0$ and (7) holds at the original tariff vector.

For foreign country $*1$, we get

$$dW^{*1} = \frac{dW^{*1}}{d\tau}(-\epsilon) + \frac{\partial W^{*1}}{\partial \tau^{*1}}(-\frac{dW^{*1}}{d\tau}/\frac{\partial W^{*1}}{\partial \tau^{*1}})(-\epsilon) + \frac{\partial W^{*1}}{\partial \tau^2}\sigma + \frac{\partial W^{*1}}{\partial \tau^{*2}}(-\frac{\partial W^{*1}}{\partial \tau^2}/\frac{\partial W^{*1}}{\partial \tau^{*2}})\sigma = 0,$$

which confirms that the perturbation maintains the welfare of foreign country $*1$.

Finally, for foreign country $*2$, we get

$$\begin{aligned}dW^{*2} &= \frac{dW^{*2}}{d\tau}(-\epsilon) + \frac{\partial W^{*2}}{\partial \tau^2}\sigma + \frac{\partial W^{*2}}{\partial \tau^{*1}}(-\frac{dW^{*1}}{d\tau}/\frac{\partial W^{*1}}{\partial \tau^{*1}})(-\epsilon) + \frac{\partial W^{*2}}{\partial \tau^{*2}}(-\frac{\partial W^{*1}}{\partial \tau^2}/\frac{\partial W^{*1}}{\partial \tau^{*2}})\sigma \\ &= [\frac{dW^{*2}}{d\tau} + \frac{\partial W^{*2}}{\partial \tau^{*1}}(-\frac{dW^{*1}}{d\tau}/\frac{\partial W^{*1}}{\partial \tau^{*1}})](-\epsilon) + [\frac{\partial W^{*2}}{\partial \tau^2} - \frac{\partial W^{*1}}{\partial \tau^2}]\sigma.\end{aligned}$$

For the final expression presented here, we know from (7) that the bracketed expression in the second term is negative. Our next task is to use (7), (17) and (18) to sign the bracketed expression in the first term.

To this end, we recall the inequality in (18):

$$-\frac{\partial W^{*1}}{\partial \tau^{*1}}/\frac{dW^{*1}}{d\tau} > -\frac{\partial W^{*2}}{\partial \tau^{*1}}/\frac{dW^{*2}}{d\tau}.$$

Using (7) and (17), we can re-write this inequality as

$$\frac{dW^{*2}}{d\tau} + \frac{\partial W^{*2}}{\partial \tau^{*1}}(-\frac{dW^{*1}}{d\tau}/\frac{\partial W^{*1}}{\partial \tau^{*1}}) < 0, \tag{23}$$

indicating that the bracketed expression in the first term to which we refer above is also negative.

It now follows that $dW^{*2} > 0$ if and only if

$$\epsilon > \{\frac{\frac{\partial W^{*2}}{\partial \tau^2} - \frac{\partial W^{*1}}{\partial \tau^2}}{\frac{dW^{*2}}{d\tau} + \frac{\partial W^{*2}}{\partial \tau^{*1}}(-\frac{dW^{*1}}{d\tau}/\frac{\partial W^{*1}}{\partial \tau^{*1}})}\}\sigma. \tag{24}$$

Thus, since the expression in braces is positive, $dW^{*2} > 0$ holds if the second-step perturbation parameterized by σ is small relative to the first-step perturbation parameterized by ϵ.

Consider now the direction of the tariff changes. Given (7) and (17), it is straightforward to see from (19)-(22) that all tariffs are reduced if $\epsilon > \sigma$. But using (7), (17) and (23), it is also straightforward to show that the expression in braces in (24) is greater than 1, and hence $\epsilon > \sigma$ is

assured by (24).

A final point is that we can ensure welfare gains for all countries, including foreign country $*1$, if we augment the perturbation to include an arbitrarily small increase in τ^{*1}. This augmented perturbation raises all three welfare levels while reducing all four tariffs.

We now have:

Proposition 4 *Suppose the MFN solution delivers an interior solution, $\boldsymbol{\tau}_m \equiv (\tau_m, \tau_m^{*1}, \tau_m, \overline{\tau}^{*2})$, as captured by (16). Suppose at this tariff vector that (7) holds with (17) then implied. Consider case 1 as defined by (18). For sufficiently small $\epsilon > 0$, $\sigma > 0$ satisfying (24), the perturbation defined in (19)-(22) when augmented with an arbitrarily small increase in τ^{*1} has the following effects: all tariffs are reduced, and all countries enjoy welfare gains.*

Case 2 The second case captures the possibility that, at the MFN solution and holding fixed the induced MFN world price $\widetilde{p}^w(\tau_m, \tau_m^{*1}, \overline{\tau}^{*2})$, the home country would prefer less trade whereas foreign country $*1$ would prefer more trade.

We begin with a formal expression of the conditions that define our second case:

$$-\frac{\partial W}{\partial \tau^{*1}} \bigg/ \frac{dW}{d\tau} = -\frac{\partial W^{*1}}{\partial \tau^{*1}} \bigg/ \frac{dW^{*1}}{d\tau} < -\frac{\partial W^{*2}}{\partial \tau^{*1}} \bigg/ \frac{dW^{*2}}{d\tau}, \tag{25}$$

where the expressions are evaluated at the MFN solution and where total derivatives again reflect the fact that $\tau \equiv \tau^1 = \tau^2$ in the MFN setting. As before, the equality in (25) captures a tangency between the iso-welfare curves of the home country and foreign country $*1$ in a graph with τ on the y axis and τ^{*1} on the x axis, which we now illustrate in the left panel in Figure 4. The inequality indicates that the iso-world-price curve, which gives combinations of τ and τ^{*1} that preserve $\widetilde{p}^w(\tau, \tau^{*1}, \tau^{*2})$ at the value obtained at the MFN solution, is now steeper, as Figure 4 depicts. This ensures that, at the prevailing world price, the home country desires less trade whereas foreign country $*1$ prefers more trade. Finally, since foreign country $*2$ is impacted by changes in τ and τ^{*1} only insofar as those changes impact the world price, we may again equivalently express the slope of the iso-world-price curve in terms of the slope of iso-welfare curve for foreign country $*2$.

The idea of the perturbation is related to that in the first case, although a different set of tariff movements is now required. In step 1 of the perturbation, the home country and foreign country $*1$ now raise τ and τ^{*1} so as to maintain W^{*1}. This step is illustrated in Figure 4 with the movement from point A to point B. Under this second case, this change necessitates a terms-of-trade improvement for foreign country $*1$ (i.e., an increase in $\widetilde{p}^w$) as compensation for the reduced trade volume. Due to the tangency condition (16), the home country again experiences only a second-order loss from the first-step adjustment. Foreign country $*2$, however, enjoys a first-order welfare gain from the first-step adjustment, since foreign country $*2$ has the same terms-of-trade as foreign country $*1$ and thus enjoys a terms-of-trade gain from the first-step adjustment. As before, in the second step of the perturbation, we utilize the full tariff space and raise τ^2 while

lowering τ^{*2}, again so as to maintain W^{*1}. We illustrate this step in Figure 4 with the movement from point C to point D. This second change generates a first-order gain for the home country and a first-order loss for foreign country $*2$. If the second-step adjustment is sufficiently small in magnitude, however, then the home country and foreign country $*2$ enjoy first-order gains overall. From here, we could allow a very small increase in the tariff of foreign country $*1$ to ensure that all countries gain from the perturbation.

One difference between the first and second cases is that not all tariffs move in the same direction in the perturbation that we use for the second case. In particular, in the perturbation that we use for the second case, τ^1, τ^2 and τ^{*1} rise while τ^{*2} falls. The basic perturbation is otherwise similar to that undertaken in our analysis of Case 1. The complete formal argument now follows.

Starting at a tariff vector that satisfies (16) and (25), and under the assumption given in (7) with (17) then implied, we now consider the following perturbation:

$$d\tau^1 = \epsilon \tag{26}$$

$$d\tau^2 = \epsilon + \sigma \tag{27}$$

$$d\tau^{*1} = (-\frac{dW}{d\tau} / \frac{\partial W}{\partial \tau^{*1}})\epsilon = (-\frac{dW^{*1}}{d\tau} / \frac{\partial W^{*1}}{\partial \tau^{*1}})\epsilon \tag{28}$$

$$d\tau^{*2} = (-\frac{\partial W^{*1}}{\partial \tau^2} / \frac{\partial W^{*1}}{\partial \tau^{*2}})\sigma \tag{29}$$

where the equality in (28) follows from the bilateral efficiency condition (16) which the starting tariffs are assumed to satisfy, and where $\epsilon > 0$ and $\sigma > 0$ are both small. We give a further condition below concerning the relative magnitudes of ϵ and σ.

Just as before, we can now compute the welfare differentials. For the home country, we get

$$dW = \frac{dW}{d\tau}\epsilon + \frac{\partial W}{\partial \tau^2}\sigma + \frac{\partial W}{\partial \tau^{*1}}(-\frac{dW}{d\tau} / \frac{\partial W}{\partial \tau^{*1}})\epsilon + \frac{\partial W}{\partial \tau^{*2}}(-\frac{\partial W^{*1}}{\partial \tau^2} / \frac{\partial W^{*1}}{\partial \tau^{*2}})\sigma.$$

Thus,

$$dW = [\frac{\partial W}{\partial \tau^2} + \frac{\partial W}{\partial \tau^{*2}}(-\frac{\partial W^{*1}}{\partial \tau^2} / \frac{\partial W^{*1}}{\partial \tau^{*2}})]\sigma > 0,$$

where the inequality follows since $\sigma > 0$ and (7) holds at the original tariff vector.

For foreign country $*1$, we get

$$dW^{*1} = \frac{dW^{*1}}{d\tau}\epsilon + \frac{\partial W^{*1}}{\partial \tau^{*1}}(-\frac{dW^{*1}}{d\tau} / \frac{\partial W^{*1}}{\partial \tau^{*1}})\epsilon + \frac{\partial W^{*1}}{\partial \tau^2}\sigma + \frac{\partial W^{*1}}{\partial \tau^{*2}}(-\frac{\partial W^{*1}}{\partial \tau^2} / \frac{\partial W^{*1}}{\partial \tau^{*2}})\sigma = 0,$$

which confirms that the perturbation maintains the welfare of foreign country $*1$.

Finally, for foreign country $*2$, we get

$$
\begin{aligned}
dW^{*2} &= \frac{dW^{*2}}{d\tau}\epsilon + \frac{\partial W^{*2}}{\partial \tau^2}\sigma + \frac{\partial W^{*2}}{\partial \tau^{*1}}(-\frac{dW^{*1}}{d\tau}/\frac{\partial W^{*1}}{\partial \tau^{*1}})\epsilon + \frac{\partial W^{*2}}{\partial \tau^{*2}}(-\frac{\partial W^{*1}}{\partial \tau^2}/\frac{\partial W^{*1}}{\partial \tau^{*2}})\sigma \\
&= [\frac{dW^{*2}}{d\tau} + \frac{\partial W^{*2}}{\partial \tau^{*1}}(-\frac{dW^{*1}}{d\tau}/\frac{\partial W^{*1}}{\partial \tau^{*1}})]\epsilon + [\frac{\partial W^{*2}}{\partial \tau^2} - \frac{\partial W^{*1}}{\partial \tau^2}]\sigma.
\end{aligned}
$$

For the final expression presented here, we know from (7) that the bracketed expression in the second term is negative. Our next task is to use (7), (17) and (25) to sign the bracketed expression in the first term.

To this end, we recall the inequality in (25):

$$
-\frac{\partial W^{*1}}{\partial \tau^{*1}}/\frac{dW^{*1}}{d\tau} < -\frac{\partial W^{*2}}{\partial \tau^{*1}}/\frac{dW^{*2}}{d\tau}.
$$

Using (7) and (17), we can re-write this inequality as

$$
\frac{dW^{*2}}{d\tau} + \frac{\partial W^{*2}}{\partial \tau^{*1}}(-\frac{dW^{*1}}{d\tau}/\frac{\partial W^{*1}}{\partial \tau^{*1}}) > 0,
$$

indicating that the bracketed expression in the first term to which we refer above is positive.

It now follows that $dW^{*2} > 0$ if and only if

$$
\epsilon > \{\frac{-[\frac{\partial W^{*2}}{\partial \tau^2} - \frac{\partial W^{*1}}{\partial \tau^2}]}{\frac{dW^{*2}}{d\tau} + \frac{\partial W^{*2}}{\partial \tau^{*1}}(-\frac{dW^{*1}}{d\tau}/\frac{\partial W^{*1}}{\partial \tau^{*1}})}\}\sigma. \tag{30}
$$

Thus, since the expression in braces is positive, $dW^{*2} > 0$ holds if the second-step perturbation parameterized by σ is small relative to the first-step perturbation parameterized by ϵ.

Consider now the direction of the tariff changes. Given (7) and (17), it is straightforward to see from (26)-(29) that all tariffs are increased except for τ^{*2} which is reduced.

A final point is that we can ensure welfare gains for all countries, including foreign country $*1$, if we augment the perturbation to include an arbitrarily small increase in τ^{*1}. This augmented perturbation raises all three welfare levels.

We now have:

Proposition 5 *Suppose the MFN solution delivers an interior solution, $\boldsymbol{\tau}_m \equiv (\tau_m, \tau_m^{*1}, \tau_m, \overline{\tau}^{*2})$, as captured by (16). Suppose at this tariff vector that (7) holds with (17) then implied. Consider case 2 as defined by (25). For sufficiently small $\epsilon > 0$, $\sigma > 0$ satisfying (30), the perturbation defined in (26)-(29) when augmented with an arbitrarily small increase in τ^{*1} has the following effects: all tariffs are raised except for τ^{*2} which is reduced, and all countries enjoy welfare gains.*

Case 3 A final case occurs when, at the MFN solution and holding fixed the induced MFN world price $\widetilde{p}^w(\tau_m, \tau_m^{*1}, \overline{\tau}^{*2})$, the home country and foreign country $*1$ each achieve their preferred levels of trade.

Formally, this case arises when

$$-\frac{\partial W}{\partial \tau^{*1}}\Big/\frac{dW}{d\tau} = -\frac{\partial W^{*1}}{\partial \tau^{*1}}\Big/\frac{dW^{*1}}{d\tau} = -\frac{\partial W^{*2}}{\partial \tau^{*1}}\Big/\frac{dW^{*2}}{d\tau}, \tag{31}$$

where the expressions are evaluated at the MFN solution and where total derivatives again reflect the fact that $\tau \equiv \tau^1 = \tau^2$ in the MFN setting. In the case captured by (31), for a graph with τ on the y axis and τ^{*1} on the x axis, the home country iso-welfare curve, the iso-welfare curve of foreign country $*1$ and the iso-world-price locus are all tangent at the tariff pair corresponding to the MFN solution, as depicted in Figure 5.

One particular example of this kind occurs when the tariff vector corresponds to the MFN politically optimal tariff vector (so that τ^{*2} is set at a particular value, too).[24] Bagwell and Staiger (1999) show that this tariff vector is efficient within the full set of discriminatory tariffs. Thus, as a general matter, it is not always possible in this case to engineer a perturbation that increases welfare for each of the three countries.

We conclude this section with a brief discussion concerning the necessary features of perturbations defined in the full tariff space that generate Pareto gains when we start at the MFN solution. Recall that in Section 5 we provided such characterizations starting at the Horn-Wolinsky solution for a setting in which the MFN requirement is not imposed. Once the MFN requirement is imposed, the necessary features of Pareto-improving perturbations are more challenging to characterize. Whether or not the MFN requirement is imposed on the solution concept, we have an unlimited space of tariff perturbations to try and restrict. In the absence of an MFN requirement, however, the home country negotiates with each foreign country, and so the Horn-Wolinsky solution concept generates tangency restrictions for each bilateral relationship. By contrast, the MFN solution doesn't generate a similar tangency restriction between the home-country MFN tariff and the tariff of the non principal supplier country (i.e., foreign country $*2$). We lose some leverage for this reason, making it more difficult to restrict the set of Pareto-improving perturbations.[25]

7   The MFN-Constrained Efficiency Frontier

In the previous section we considered the efficiency properties of the MFN solution when evaluated relative to the full space of tariff policies $\boldsymbol{\tau} \equiv (\tau^1, \tau^{*1}, \tau^2, \tau^{*2})$. We now assess the efficiency properties of this solution when judged by the MFN-constrained efficiency frontier. To this end,

---

[24]In particular, in the MFN politically optimal tariff vector, τ^{*2} is set at the value that generates tangencies as depicted in Figure 5, once all appearances of "$*1$" and "$*2$" in that figure and on the x-axis are interchanged.

[25]Starting at the MFN solution and allowing for discriminatory perturbations, suppose that, in a graph with τ^2 and τ^{*2} on the axes, there is an upward lens (not a tangency) formed by the home-country and foreign-country $*2$ iso-welfare curves. Due to the first-order welfare gains that higher (within-lens) values for τ^2 and τ^{*2} generate for the home country and foreign country $*2$, it is then a simple matter to construct a Pareto-improving perturbation in which all four tariffs are increased. Our sufficiency propositions presented in this section, by contrast, are more robust in that they do not impose an assumption about the initial direction of any tariff lens in the relationship between the home country and foreign country $*2$. For general settings, they identify specific perturbations that generate Pareto gains while lowering all tariffs (Proposition 4) or at least some tariff (Proposition 5).

we ask the following question: Under the assumption given in (7) with (17) then implied, what are the efficiency properties of the MFN solution $\boldsymbol{\tau}_m \equiv (\tau_m, \tau_m^{*1}, \tau_m, \overline{\tau}^{*2})$ as captured by (16) when efficiency is defined relative to the space of MFN tariff policies, $(\tau, \tau^{*1}, \tau, \tau^{*2})$, where $\tau \equiv \tau^1 = \tau^2$ is the home country's MFN tariff?

To answer this question, we now assume that the country welfare functions W and W^{*i} for $i = 1, 2$ satisfy standard regularity conditions so that each point on the MFN-constrained efficiency frontier is interior in that $(\tau, \tau^{*1}, \tau, \tau^{*2}) \in (\underline{\tau}, \overline{\tau})^4$, and we begin by exploiting an implication of the characterization of the MFN-constrained efficiency frontier derived in Bagwell and Staiger (2005). According to their Proposition 7, at an MFN-efficient vector of tariffs, either

$$-\frac{\partial W^{*1}}{\partial \tau^{*1}} \bigg/ \frac{dW^{*1}}{d\tau} < -\frac{\partial W}{\partial \tau^{*1}} \bigg/ \frac{dW}{d\tau} < -\frac{\partial \widetilde{p}^w}{\partial \tau^{*1}} \bigg/ \frac{\partial \widetilde{p}^w}{\partial \tau}; \tag{32}$$

$$-\frac{\partial \widetilde{p}^w}{\partial \tau^{*1}} \bigg/ \frac{\partial \widetilde{p}^w}{\partial \tau} < -\frac{\partial W}{\partial \tau^{*1}} \bigg/ \frac{dW}{d\tau} < -\frac{\partial W^{*1}}{\partial \tau^{*1}} \bigg/ \frac{dW^{*1}}{d\tau}; \text{ or} \tag{33}$$

$$-\frac{\partial \widetilde{p}^w}{\partial \tau^{*1}} \bigg/ \frac{\partial \widetilde{p}^w}{\partial \tau} = -\frac{\partial W}{\partial \tau^{*1}} \bigg/ \frac{dW}{d\tau} \underset{<}{\overset{>}{=}} -\frac{\partial W^{*1}}{\partial \tau^{*1}} \bigg/ \frac{dW^{*1}}{d\tau}, \tag{34}$$

where total derivatives reflect that $\tau \equiv \tau^1 = \tau^2$ in the MFN setting. In writing these expressions, we use the fact that foreign country $*2$ is impacted by changes in τ and τ^{*1} only insofar as those changes impact the world price, so that in (τ, τ^{*1}) space we may equivalently express the slope of the iso-welfare curve for foreign country $*2$ as the slope of the iso-world-price curve.

At points on the MFN-constrained efficiency frontier satisfying (32), the home country desires less trade whereas foreign country $*1$ prefers more trade at the prevailing world price. By contrast, at points on the MFN-constrained efficiency frontier satisfying (33), the home country desires more trade whereas foreign country $*1$ prefers less trade at the prevailing world price. And finally, at points on the MFN-constrained efficiency frontier satisfying (34), the home country achieves its desired trade volume whereas foreign country $*1$ may desire more trade, desire less trade, or achieve its desired trade volume as well at the prevailing world price. Notice that under condition (32), the home country's iso-welfare curve is steeper than that of foreign country $*1$, indicating a downward lens between the home country and foreign country $*1$ in (τ, τ^{*1}) space, while under condition (33) the home country's iso-welfare curve is flatter than that of foreign country $*1$, indicating an upward lens between the home country and foreign country $*1$. Finally, under condition (34), there can be a downward lens, an upward lens, or no lens at all between the home country and foreign country $*1$, with this last possibility corresponding to the MFN politically optimal tariffs.

Consider now three cases. A first case is where τ^{*2} is fixed at a level $\overline{\tau}^{*2}$ that is consistent with a point on the MFN-constrained efficiency frontier that satisfies (32). In this case, as we have observed, there is a downward lens between the home country and foreign country $*1$. But the MFN solution is characterized by a point of tangency between the home-country and foreign-country $*1$ iso-welfare curves as described by (16), so it follows that for this case there exists a range of values for the initial tariff vector $\boldsymbol{\tau}_0$ and the bargaining parameter α such that the

MFN solution corresponds to a point of tangency within this lens and therefore entails greater liberalization between the home country and foreign country $*1$ than is MFN-constrained efficient given $\tau^{*2} \equiv \overline{\tau}^{*2}$. A second case is where τ^{*2} is fixed at a level $\overline{\tau}^{*2}$ that is consistent with a point on the MFN-constrained efficiency frontier that satisfies (33). In this case, as we have observed, there is an upward lens between the home country and foreign country $*1$. With the MFN solution described by (16), it then follows that for this case there exists a range of values for the initial tariff vector $\boldsymbol{\tau}_0$ and the bargaining parameter α such that the MFN solution corresponds to a point of tangency within this lens and therefore entails less liberalization between the home country and foreign country $*1$ than is MFN-constrained efficient given $\tau^{*2} \equiv \overline{\tau}^{*2}$. A third case is where τ^{*2} is fixed at a level $\overline{\tau}^{*2}$ that is consistent with a point on the MFN-constrained efficiency frontier that satisfies (34). As we have observed, in this case there can be a downward lens, an upward lens, or no lens at all between the home country and foreign country $*1$, with this last possibility corresponding to the MFN politically optimal tariffs. It then follows that for this case the MFN solution may entail less liberalization or more liberalization between the home country and foreign country $*1$ than is MFN-constrained efficient given $\tau^{*2} \equiv \overline{\tau}^{*2}$; or the MFN solution could achieve the MFN-efficient political optimum.[26]

This last possibility is of some special interest. Evidently, if τ^{*2} is fixed at its MFN politically optimal level, there exists some initial tariff vector $\boldsymbol{\tau}_0$ and some bargaining parameter α such that the MFN solution for the bargain between the home country and foreign country $*1$ would deliver countries to the MFN-efficient political optimum. Of course, this is an extremely special set of conditions. More generally, we can conclude that, with the exception of this knife-edge case, the MFN solution will be inefficient, and it can lead to either too much liberalization or too little liberalization relative to the MFN-constrained efficiency frontier.

We summarize with:

Proposition 6 *Suppose the MFN solution delivers an interior solution, $\boldsymbol{\tau}_m \equiv (\tau_m, \tau_m^{*1}, \tau_m, \overline{\tau}^{*2})$, as captured by (16). Suppose at this tariff vector that (7) holds with (17) then implied. And finally, suppose that each point on the MFN-constrained efficiency frontier is interior. Then the MFN solution is generically inefficient relative to the MFN-constrained efficiency frontier, and may lead to either too little liberalization or too much liberalization relative to MFN-constrained efficient levels.*

In light of the positive externality that the home country's MFN tariff liberalization imparts on foreign country $*2$, it may seem surprising that the MFN solution could ever lead to too much liberalization relative to MFN-constrained efficient tariff levels.[27] But recalling that, beginning from MFN-efficient tariff levels, the home country and foreign country $*1$ can only gain in their

---

[26]As shown by Bagwell and Staiger (1999) and noted in the previous section, the MFN political optimum is efficient relative to the full space of tariffs. It then follows immediately that the MFN political optimum is also MFN-constrained efficient.

[27]Recall that, under (7) and the implied (17), an MFN tariff cut by the home country generates a positive externality (i.e., a welfare gain) for each foreign country.

bilateral bargain if they worsen foreign country $*2$'s – and hence under MFN, also foreign country $*1$'s – terms of trade, the bilateral bargain must move foreign country $*1$'s trade volume in the direction that foreign country $*1$ would desire at a fixed terms of trade; and if at the relevant point on the MFN-constrained efficiency frontier foreign country $*1$ desires more trade volume at a fixed terms of trade, then according to the MFN solution the bilateral bargain between the home country and foreign country $*1$ will lead these countries to reduce their tariffs and engage in too much liberalization relative to MFN-constrained efficient tariff levels.[28]

8 Endowment Economy with Cobb-Douglas Preferences

In this section, we analyze the three-country, two-good general equilibrium model developed above under the further assumptions of an endowment economy with Cobb-Douglas (CD) preferences wherein both products receive equal weight. We assume that each government maximizes the national economic welfare (real national income) of its country, which is captured as the indirect utility of the representative agent. Our numerical analysis of this particular representation enables us to illustrate concretely and further develop many themes raised above. We also verify the existence of the interior Horn-Wolinsky solution for this representation. In an online appendix containing supplementary material, we explore this representation in detail and derive the analytic welfare expressions reported below.

8.1 The Model

We assume now that the utility for a representative agent in the home country is given by $U(x_c, y_c)$ when x_c units of good x and y_c units of good y are consumed. Given our specification of CD preferences with both products weighted equally, we have that

$$U(x_c, y_c) = x_c \cdot y_c. \tag{35}$$

Similarly, the utility function for a representative consumer in foreign country $*j$, where $j = 1, 2$, is given by

$$U^{*j}(x_c^{*j}, y_c^{*j}) = x_c^{*j} \cdot y_c^{*j}, \tag{36}$$

---

[28]An interesting question is whether more can be said about the conditions under which each of the possibilities described in Proposition 6 might arise. While we can't offer a complete answer to this question here, if we think of the tariff of foreign country $*2$ as being fixed on its reaction curve, and if governments are sufficiently close to national income maximizers, then it is straightforward to establish that, at the relevant point on the MFN-constrained efficiency frontier, foreign country $*1$ desires more trade volume at the fixed terms of trade while the home country desires less trade volume at the fixed terms of trade; and hence, under these conditions, a downward lens is associated with this point. We cannot conclude that the MFN solution entails too much liberalization even in this case, however, since the solution may fall outside this lens depending on the bargaining parameter if the threat point corresponds to a distinct tariff pair, such as the best-response tariff pair for the home country and foreign country *1. We explore these issues more deeply in Section 8, where we examine an endowment economy with Cobb-Douglas preferences when each government maximizes the indirect utility of the representative agent in its country.

when x_c^{*j} units of good x and y_c^{*j} units of good y are consumed. The corresponding utility-maximizing consumption levels for the home-country representative agent are

$$x_c(p, I) = I/(2p) \text{ and } y_c(p, I) = I/2, \tag{37}$$

where $p \equiv p_x/p_y$ and I is home country income (inclusive of tariff revenue) expressed in local units of good y. Similarly, in foreign country $*j$, where $j = 1, 2$, the utility-maximizing consumption levels are given by

$$x_c^{*j}(p^{*j}, I^{*j}) = I^{*j}/(2p^{*j}) \text{ and } y_c^{*j}(p^{*j}, I^{*j}) = I^{*j}/2, \tag{38}$$

where $p^{*j} \equiv p_x^{*j}/p_y^{*j}$ and I^{*j} is foreign country $*j$ income (inclusive of tariff revenue) expressed in local units of good y.

The supply side of the endowment economy is straightforward to describe. Let Q_x^{*j} and Q_y^{*j} denote the endowments of good x and y, respectively, in foreign country $*j$, where $j = 1, 2$. Similarly, the endowments of good x and y, respectively, in the home country are denoted as Q_x and Q_y. As in previous sections, we assume that the endowments are such that the home country exports (imports) good y (x), while each foreign country $*j$ exports (imports) good x (y). We define explicit symmetric and asymmetric specifications for endowments below.

In our online appendix, we solve the model and characterize welfare functions. Here, we simply report the relevant findings. Imposing the price relationships $p = \tau^j p^{wj}$ and $p^{*j} = p^{wj}/\tau^{*j}$, we derive expressions in the online appendix for tariff revenues, incomes, trade volumes and indirect utilities. We focus here on the expressions for indirect utilities. The indirect utility function for the representative agent in foreign country $*j$ is

$$V^{*j}(p^{*j}, I^{*j}(p^{*j}, p^{wj})) = \left(\frac{p^{wj}Q_x^{*j} + Q_y^{*j}}{p^{wj} + p^{*j}}\right)^2 p^{*j}, \tag{39}$$

where the notation reflects that I^{*j} is implied once p^{*j} and p^{wj} are determined. Similarly, the indirect utility function for the representation agent in the home country is

$$V(p, I(p, T)) = \left(\frac{Q_x T + Q_y}{p + T}\right)^2 p, \tag{40}$$

where we recall that $1/T$ is the home country's multilateral terms of trade with T having the functional dependence, $T = T(p^{*1}, p^{*2}, p^{w1}, p^{w2})$, as discussed in Section 2. As the notation in (40) reflects, income in the home country is implied once p and T are determined.

Given tariffs and for $j = 1, 2$, once the market clearing world prices are determined, the local prices p and p^{*j} are implied via the relationships $p = \tau^j p^{wj}$ and $p^{*j} = p^{wj}/\tau^{*j}$. Since $T = T(p^{*1}, p^{*2}, p^{w1}, p^{w2})$, it then follows that the home country's multilateral terms of trade, $1/T$, is also determined given tariffs, once the market clearing world prices are determined. In line with our discussion in Section 2 and as we formally show in the online appendix, the two market clearing world prices in turn can be determined as functions of tariffs by using a market clearing condition

and the "linkage condition" under which $p^{w2} = \frac{\tau^1}{\tau^2} p^{w1}$. As before, we denote the market clearing world prices as $\widetilde{p}^{w1}(\boldsymbol{\tau})$ and $\widetilde{p}^{w2}(\boldsymbol{\tau})$. For the endowment economy with CD preferences, we show in the online appendix that the market clearing world prices are given as

$$\widetilde{p}^{wj}(\boldsymbol{\tau}) = \frac{Q_y + Q_y^{*j}(\frac{\tau^{*j}(1+\tau^j)}{1+\tau^{*j}}) + Q_y^{*i}(\frac{\tau^{*i}(1+\tau^i)}{1+\tau^{*i}})}{\tau^j Q_x + Q_x^{*j}(\frac{1+\tau^j}{1+\tau^{*j}}) + Q_x^{*i}(\frac{\tau^j}{\tau^i})(\frac{1+\tau^i}{1+\tau^{*i}})},$$

(41)

where $i, j = 1, 2$ and $i \neq j$.

With market clearing world prices determined, we may define the welfare function for foreign country $*j$ as

$$W^{*j}(\boldsymbol{\tau}) \equiv V^{*j}(\widetilde{p}^{wj}/\tau^{*j}, I^{*j}(\widetilde{p}^{wj}/\tau^{*j}, \widetilde{p}^{wj})),$$

where we use the relationship $p^{*j} = p^{wj}/\tau^{*j}$ and for notational ease suppress the dependency of $\widetilde{p}^{wj}$ on $\boldsymbol{\tau}$.[29] We thus have from (39) that

$$W^{*j}(\boldsymbol{\tau}) = \left(\frac{\widetilde{p}^{wj} Q_x^{*j} + Q_y^{*j}}{1 + \tau^{*j}}\right)^2 \frac{\tau^{*j}}{\widetilde{p}^{wj}},$$

(42)

where $\widetilde{p}^{wj}$ is given in (41).

To represent the welfare of the home country, we first introduce notation for the home country's multilateral terms of trade at market clearing prices. Using the relationship $p^{*j} = p^{wj}/\tau^{*j}$ for $j = 1, 2$, let

$$\widetilde{T}(\boldsymbol{\tau}) \equiv T(\widetilde{p}^{w1}/\tau^{*1}, \widetilde{p}^{w2}/\tau^{*2}, \widetilde{p}^{w1}, \widetilde{p}^{w2}).$$

(43)

We may define the welfare function for the home country as

$$W(\boldsymbol{\tau}) \equiv V(\tau^1 \widetilde{p}^{w1}, I(\tau^1 \widetilde{p}^{w1}, \widetilde{T})),$$

where we use the relationship $p = \tau^1 p^{w1}$ and (43).[30] We thus have from (40) that

$$W(\boldsymbol{\tau}) = \left(\frac{Q_x \widetilde{T} + Q_y}{\tau^1 \widetilde{p}^{w1} + \widetilde{T}}\right)^2 \tau^1 \widetilde{p}^{w1}$$

(44)

where $\widetilde{p}^{w1}$ is given by (41) and where as we show in the online appendix

$$\widetilde{T}(\boldsymbol{\tau}) = \frac{(\widetilde{A}^{*1} + \widetilde{A}^{*2})\widetilde{p}^{w1}\widetilde{p}^{w2}}{\widetilde{A}^{*1}\widetilde{p}^{w2} + \widetilde{A}^{*2}\widetilde{p}^{w1}},$$

(45)

with

$$\widetilde{A}^{*j}(\boldsymbol{\tau}) \equiv \left(Q_x^{*j}\left(\frac{\widetilde{p}^{wj}}{\tau^{*j}}\right) - Q_y^{*j}\right)\left(\frac{\tau^{*j}}{\tau^{*j} + 1}\right)$$

(46)

---

[29]Observe that, in terms of the notation presented in Section 2 and using the relationship $p^{*j} = p^{wj}/\tau^{*j}$, $w^{*j}(p^{*j}, \widetilde{p}^{wj}) \equiv V^{*j}(\widetilde{p}^{wj}/\tau^{*j}, I^{*j}(\widetilde{p}^{wj}/\tau^{*j}, \widetilde{p}^{wj}))$.

[30]Observe that, in terms of the notation presented in Section 2 and using $p = \tau^1 p^{w1}$, $w(p, T) \equiv V(\tau^1 \widetilde{p}^{w1}, I(\tau^1 \widetilde{p}^{w1}, \widetilde{T}))$, where we use as well that $T = \widetilde{T}$ at market clearing prices.

and where for notational ease we suppress the dependency in (45) of $\widetilde{A}^{*1}$ and $\widetilde{A}^{*2}$ on $\boldsymbol{\tau}$.

At this point, for the model with CD preferences and arbitrary fixed endowments, we have an explicit representation for each of the three countries of its welfare as a function of the underlying four tariffs. Note that our derivations are all made under the assumption that the four tariffs are non-prohibitive. Our next step is to specify endowment level relationships. We consider two specifications: a symmetric specification and an asymmetric specification.

In our *symmetric specification*, we parameterize endowments with a single parameter $\gamma \in (0, 1/2)$ according to

$$
\begin{aligned}
Q_x &= \gamma, \ Q_y = 1 - \gamma \\
Q_x^{*j} &= \frac{1 - \gamma}{2}, \ Q_y^{*j} = \frac{\gamma}{2}, \ j = 1, 2.
\end{aligned}
$$

Then, for $\gamma \in (0, 1/2)$, we may use (42) and (44) to represent the welfare functions for foreign country $*j$, $j = 1, 2$, and the home country, respectively, where $\widetilde{p}^{wj}$ and $\widetilde{T}$ are determined once the symmetric specification values are substituted into (41), (45) and (46). As we show in our online appendix, it is straightforward to characterize autarky prices for this specification and thus to represent the set of non-prohibitive tariffs.

In our *asymmetric specification*, we include an additional parameter $\beta \in (0, 1)$ that captures any asymmetry in size between the two foreign countries. The asymmetric specification for endowments entails two parameters, $\gamma \in (0, 1/2)$ and $\beta \in (0, 1)$, where

$$
\begin{aligned}
Q_x &= \gamma, \ Q_y = 1 - \gamma \\
Q_x^{*1} &= \beta(1 - \gamma), \ Q_y^{*1} = \beta\gamma \\
Q_x^{*2} &= (1 - \beta)(1 - \gamma), \ Q_y^{*2} = (1 - \beta)\gamma
\end{aligned}
$$

Then, for $\gamma \in (0, 1/2)$ and $\beta \in (0, 1)$, we may use (42) and (44) to represent the welfare functions for foreign country $*j$, $j = 1, 2$, and the home country, respectively, where where $\widetilde{p}^{wj}$ and $\widetilde{T}$ are determined once the asymmetric specification values are substituted into (41), (45), and (46). For the asymmetric specification as well, it is straightforward to derive autarky prices and associated conditions for non-prohibitive tariffs

8.2 Characterizations

We now provide characterizations of Nash, efficient and Horn-Wolinsky tariffs under our symmetric and asymmetric endowment specifications, respectively.

Symmetric Specification: Nash and Horn-Wolinsky Tariffs   We start with the symmetric specification. Table 1 provides numerical solutions for Nash tariffs at different values for γ. Observe that lower values for γ indicate a greater difference between the endowments of the home country and the foreign countries and are thus associated with higher trade volumes. As the table illustrates,

for the symmetric setting, the home-country tariffs are symmetric and the foreign-country tariffs are also symmetric. Symmetric tariffs are an outcome under the symmetric specification; we do not impose a symmetry restriction on the strategy space.

Table 1: Nash Equilibrium Tariff

	$\tau_1 = \tau_2$	$\tau_1^* = \tau_2^*$
$\gamma = 1e-6$	1225.83	1.99
$\gamma = 0.05$	5.26	1.66
$\gamma = 0.1$	3.54	1.50
$\gamma = 0.25$	1.92	1.24
$\gamma = 0.4$	1.27	1.08
$\gamma = 0.45$	1.13	1.04

We next consider the Horn-Wolinsky tariffs for the symmetric specification. The Horn-Wolinsky solution for the symmetric specification depends on the endowment parameter γ, the bargaining power parameter α, and the choice of a threat point for bilateral negotiations. Recall also that an *interior Horn-Wolinsky solution* must be on the bilateral efficiency frontier:

$$\frac{\partial W}{\partial \tau^{*i}} \Big/ \frac{\partial W}{\partial \tau^i} = \frac{\partial W^{*i}}{\partial \tau^{*i}} \Big/ \frac{\partial W^{*i}}{\partial \tau^i}, \ i = 1, 2. \tag{47}$$

Table 2 presents our numerical characterization of the Horn-Wolinsky solution for different scenarios. Different values for γ are again considered, and we also allow for different values of α. We also consider two possible threat points for bilateral negotiations: we allow that a failed bilateral negotiation would cause the tariffs under negotiation to revert to Nash equilibrium tariffs (NE) or to move to free trade (FT). The NE case can be motivated by a setting in which pre-negotiation tariffs are set at Nash levels, while the FT case can be motivated if we consider a situation in which bilateral negotiations are contemplated even though the initial tariffs are already at free trade. In each case, the Horn-Wolinsky solution satisfies (47); thus, Table 2 illustrates interior Horn-Wolinsky solutions under different parameter settings and threat-point cases. To see the effect of an increase in the home country's bargaining power α, we may refer to the first two parameter specifications in the table. We see that the interior Horn-Wolinsky tariff for the home country (each foreign country) is higher (lower) when α is higher, where the differences are more pronounced under the NE threat point.

We observe from Table 2 that, when the initial tariffs are already at free trade, the interior Horn-Wolinsky solution calls for further tariff cuts, resulting in import subsidies. By contrast, when the Nash threat point is used, the interior Horn-Wolinsky solution again results in tariff reductions but it may but need not hold that all negotiated tariffs entail import subsidies. For

Table 2: Horn-Wolinsky Solution to the Tariff Negotiation

		Threat Point	$\tau_1^{hw} = \tau_2^{hw}$	$\tau_1^{*hw} = \tau_2^{*hw}$
$\gamma = 0.25$	$\alpha = 0.5$	NE	0.85	0.75
		FT	0.82	0.76
$\gamma = 0.25$	$\alpha = 0.75$	NE	1.06	0.66
		FT	0.84	0.75
$\gamma = 0.4$	$\alpha = 0.5$	NE	0.922	0.895
		FT	0.919	0.897
$\gamma = 0.05$	$\alpha = 0.5$	NE	1.05	0.51
		FT	0.73	0.62

Notes: NE is for Nash Equilibrium and FT is for Free Trade.

example, when $\gamma = 0.25$ and the home country has significant bargaining power with $\alpha = 0.75$, the negotiated tariff for the foreign countries entails an import subsidy while the negotiated tariff for the home country entails a (small) import tariff. Since global free trade is one of a continuum of efficient tariff policies for this set up, the findings reported in Table 2 are consistent with the excessive liberalization results under discriminatory tariffs as reported in Propositions 1-3.

Asymmetric Specification: Nash and Horn-Wolinsky Tariffs   We now consider the asymmetric specification. Table 3 illustrates the Nash tariffs for different values of γ and β. Notice that the Nash tariffs are no longer symmetric, since the foreign countries are themselves asymmetric when $\beta \neq 1/2$. For simplicity, we consider specifications under which $\beta > 1/2$ so that foreign country $*1$ is larger than foreign country $*2$.

Table 3: Nash Equilibrium Tariff

	β	τ_1	τ_2	τ_1^*	τ_2^*
$\gamma = 0.25$	2/3	1.85	2.00	1.35	1.15
	5/6	1.79	2.12	1.51	1.07
$\gamma = 0.4$	2/3	1.26	1.30	1.12	1.05
	5/6	1.24	1.32	1.17	1.03

Table 4 then provides the corresponding interior Horn-Wolinsky solutions for different values of γ, α and β under Nash equilibrium and free-trade threat points, respectively.

Using Table 4, and in line with our observations for Table 2, we see that, starting at free trade, the interior Horn-Wolinsky solution calls for import subsidies by all countries. When the Nash

Table 4: Horn-Wolinsky Solution to the Tariff Negotiation

Threat Point	γ	β	α	τ_1	τ_2	τ_1^*	τ_2^*
	0.25	2/3	0.5	0.876	0.849	0.797	0.708
	0.25	2/3	0.75	1.056	1.067	0.715	0.619
	0.25	5/6	0.5	0.929	0.888	0.871	0.666
	0.25	5/6	0.75	1.064	1.095	0.795	0.582
NE	0.4	2/3	0.5	0.937	0.924	0.918	0.876
	0.4	2/3	0.75	1.002	1.004	0.881	0.832
	0.4	5/6	0.5	0.965	0.945	0.950	0.855
	0.4	5/6	0.75	1.011	1.019	0.918	0.812
	0.25	2/3	0.5	0.844	0.814	0.815	0.724
	0.25	2/3	0.75	0.857	0.838	0.809	0.710
	0.25	5/6	0.5	0.898	0.833	0.889	0.697
	0.25	5/6	0.75	0.904	0.856	0.888	0.681
FT	0.4	2/3	0.5	0.932	0.918	0.921	0.879
	0.4	2/3	0.75	0.937	0.929	0.919	0.872
	0.4	5/6	0.5	0.957	0.929	0.954	0.866
	0.4	5/6	0.75	0.909	0.914	0.903	0.838

Notes: NE is for Nash Equilibrium and FT is for Free Trade.

threat point is used, by contrast, the interior Horn-Wolinsky solution may entail a (small) positive import tariff for the home country, if the home country has sufficient bargaining power. The Horn-Wolinsky solution entails an import subsidy for both foreign countries for all parameters values that we consider.

MFN Solution   We now suppose that the home country negotiates only with foreign country $*1$ while satisfying the MFN rule: $\tau \equiv \tau^1 = \tau^2$. As discussed in Section 6, the negotiation between the home country and foreign country $*1$ is conducted according to the Nash bargaining solution, and we refer to the resulting solution as the MFN solution. The interior MFN solution satisfies

$$\frac{\partial W}{\partial \tau^{*1}} \Big/ \frac{dW}{d\tau} = \frac{\partial W^{*1}}{\partial \tau^{*1}} \Big/ \frac{dW^{*i}}{d\tau}. \tag{48}$$

In line with our work in Section 7, we consider here the relationship between the MFN solution and the MFN-constrained efficiency frontier.

Since the home country negotiates with foreign country $*1$, the natural assumption is that foreign country $*1$ is the principal supplier to the home country. We thus use the asymmetric specification with the restriction that foreign country $*1$ has a bigger export-good endowment than does foreign country $*2$. We hold foreign country $*2$'s tariff fixed at its Nash level, $\overline{\tau}^{*2} = \tau^{*2N}$. The MFN solution is thus represented by the vector $\boldsymbol{\tau}_m = (\tau_m, \tau_m^{*1}, \tau_m, \tau^{*2N})$. Given that governments maximize national economic welfare in our endowment economy with CD preferences, efficiency

requires that local relative prices are equal across countries; thus, the MFN-constrained efficiency frontier when $\tau^{*2} = \tau^{*2N}$ requires that $\tau \cdot \tau^{*2N} = 1$ and $\tau \cdot \tau^{*1} = 1$. It follows that there is a unique tariff vector that is MFN-constrained efficient. Accordingly, we define the MFN-constrained efficient tariff vector when $\tau^{*2} = \tau^{*2N}$ by the vector $(\tau^E(\tau^{*2N}), \tau^{*1E}(\tau^{*2N}), \tau^E(\tau^{*2N}), \tau^{*2N})$ where $\tau^E(\tau^{*2N}) = 1/\tau^{*2N}$ and $\tau^{*1E}(\tau^{*2N}) = \tau^{*2N}$. Our findings in Proposition 6 suggests both the possibility of too much liberalization and too little liberalization relative to the MFN-constrained efficiency frontier. With $\overline{\tau}^{*2}$ set equal to τ^{*2N}, the possibility of too much liberalization corresponds to $\tau_m < \tau^E(\tau^{*2N})$ and $\tau_m^{*1} < \tau^{*1E}(\tau^{*2N})$ while the possibility of too little liberalization corresponds to $\tau_m > \tau^E(\tau^{*2N})$ and $\tau_m^{*1} > \tau^{*1E}(\tau^{*2N})$.

To explore these possibilities, we introduce the parameter λ to represent the size of the home country relative to the rest of the world. To this end, we adjust the home endowment levels in the asymmetric specification so that they are now

$$Q_x = \lambda \cdot \gamma, \ Q_y = \lambda \cdot (1 - \gamma),$$

and we leave the foreign endowment levels unaltered from those given in the asymmetric specification. When $\lambda = 1$, the size of the home country is thus equal to the combined size of the two foreign countries, and the home country is smaller than the combined size of the foreign countries when $\lambda < 1$.

We illustrate our findings with two figures. In Figure 6, we set $\gamma = 0.1$, $\beta = 0.6$, and $\lambda = 0.03$. The horizontal axis represents the MFN tariff τ of the home country, and the vertical axis captures the tariff τ^{*1} of foreign country $*1$. Indifference (ID) curves are blue for the home country and red for foreign country $*1$. The Nash equilibrium is depicted in the Northeast region of the figure as a blue dot, where the blue (red) indifference curve has zero (infinite) slope. Taking the Nash point as the threat point for the Nash bargaining solution, we represent the bargaining locus defined by (48) by the black curve. Along this locus, the indifference curves of the home country and foreign country $*1$ are thus tangent. The contract curve is then defined by the portion of the black curve that falls between the Nash indifference curves for the home country and foreign country $*1$. Figure 6 illustrates the tangency associated with the MFN solution for the case where $\alpha = 0.5$. Other points along the contract curve correspond to MFN solutions under other values for α; in the figure, we also identify the solutions when $\alpha = 0.3$ and $\alpha = 0.05$. The pink curve is defined by $\tau \cdot \tau^{*1} = 1$, and the MFN-efficient point is then represented on this curve by the pink dot. For the parameter specification in Figure 6, we observe that the entire contract curve lies to the Southwest of the pink curve. This means that, regardless of the value of the bargaining power parameter α, the MFN solution for this specification entails too much trade liberalization.

In Figure 7, we explore an alternative parameter specification, setting $\gamma = 0.2$, $\beta = 0.9$, and $\lambda = 0.9$. The curves in Figure 7 have the same interpretation as those in Figure 6, as just described. For Figure 7, notice that the black and pink curves intersect over the region between the Nash indifference curves. Hence, for some bargaining power parameters α we get too much trade liberalization again, but for other values of α we now get too little trade liberalization.

Finally, we note that our findings as captured in Figures 6 and 7 are consistent with our results in Proposition 6, which establish that the MFN solution is generically inefficient relative to the MFN-constrained efficiency frontier, and may lead to either too little liberalization or too much liberalization relative to MFN-constrained efficient levels.

9   Discussion

In the preceding sections, we consider the efficiency properties of bilateral tariff bargaining both when tariffs can be discriminatory and when they must conform to the MFN rule. For the case of discriminatory tariffs, we use the Horn-Wolinsky solution concept to characterize the efficiency properties of simultaneous bilateral tariff bargaining. In the case where negotiations must satisfy the MFN rule, we characterize the efficiency properties of the Nash bargaining solution between the home country and foreign country $*1$ while holding the tariff of foreign country $*2$ fixed. Our representation of tariff bargaining in the presence of an MFN rule is motivated by the broad features of the GATT/WTO principal-supplier rule, but other representations are also of interest. In this section, we consider two additional features of GATT/WTO bargaining from which our analysis above has abstracted, and we discuss the potential importance of and also the challenges in extending the Horn-Wolinsky solution to environments that incorporate these features.

MFN bargaining beyond the principal supplier   In our analysis of the efficiency properties of bilateral tariff bargaining in the presence of the MFN rule, we assume that the home country only bargains with its principal supplier. Analyzing the detailed United States bargaining records of the 1950-51 GATT Torquay Round, Bagwell, Staiger and Yurukoglu (2017) confirm that engaging a single exporting country in negotiations over an MFN import tariff is the modal US behavior. However, Bagwell, Staiger and Yurukoglu also report that there are significant numbers of US bargains on a given tariff that involve more than one exporting country. These observations suggest that it could also be of interest to use the Horn-Wolinsky solution concept to characterize the efficiency properties of simultaneous bilateral bargaining under the MFN rule in our two-good model. Here we briefly comment on the challenges involved in extending the Horn-Wolinsky solution concept to encompass these broader bargaining possibilities.

The logic of the Horn-Wolinsky solution concept extends in straightforward fashion to settings with simultaneous bilateral bargaining under the MFN rule when each country imports multiple goods if each country negotiates its tariff for any given import good only with a single principal supplier of that good.[31] Some conceptual considerations arise in applying the Horn-Wolinsky solution, however, if a country negotiates its MFN tariff on a given good simultaneously with multiple partners. In particular, for the two-good model that we examine, if the home country simultaneously negotiates with both foreign countries over the MFN tariff for its import good, then the following questions must be addressed: Which negotiated MFN tariff for the home country is

---

[31]We may understand the MFN solution for our two-good model as a limit point of this approach in which the possibility of simultaneous bargaining is completely eliminated, since each country then has only one import good.

ultimately applied, and do all participants understand the process through which this determination is made at the time of their respective negotiations?

In principle, the Horn-Wolinsky solution concept could be applied in our two-good model to characterize simultaneous bilateral negotiations under MFN by introducing to the model the notion of MFN tariff *bindings* – legal commitments to a maximum applied tariff level – and allowing the home country to negotiate different binding levels on its import tariff with its different bargaining partners, with the understanding that its applied tariff cannot exceed the minimum binding level over the set of bindings it agrees to in its bargains. In this environment, an issue of multiplicity of Horn-Wolinsky equilibria arises, however.

To see the issue, let us denote the binding on its tariff τ that the home country negotiates with foreign country $*1$ by τ_{b1} and the binding on its tariff τ that the home country negotiates with foreign country $*2$ by τ_{b2}, with the understanding among all participants that the home country's tariff will then be applied at $\tau = \min[\tau_{b1}, \tau_{b2}]$ under the assumption that these binding levels are positioned below the home country's best-response tariff. Suppose now that the resulting home-country applied MFN tariff τ is such that, for each bilateral country pair, mutual gains could be enjoyed with appropriate upward movements in the home country's applied MFN tariff and the tariff of the foreign country participating in that bilateral (i.e., suppose each bilateral is characterized by an upward lens). A given bilateral country pair can engineer an increase in the home country's applied MFN tariff, however, only if the home country's minimum negotiated tariff binding is increased; hence, if $\tau_{b1} = \tau_{b2}$, then the binding negotiated in one bilateral pair would render ineffective any increase in the binding negotiated in the other bilateral pair, and for this reason no bilateral country pair would be able to engineer an upward movement in the home country's applied MFN tariff. Consequently, for any initial set of applied MFN tariffs $(\tau, \tau^{*1}, \tau^{*2})$ satisfying the "upward-lens" requirements described above, if $\tau_{b1} = \tau_{b2} = \tau$, then (τ_{b1}, τ_{b2}) corresponds to a Horn-Wolinsky equilibrium when negotiations are over MFN tariff bindings. To usefully apply the Horn-Wolinsky solution concept under the MFN rule while allowing the home country to negotiate its MFN tariff on a given good simultaneously with multiple partners, this multiplicity of equilibria will have to be addressed in some way, an issue that we leave for future research.

Article XXVIII renegotiation   In our analysis of bilateral tariff bargaining, we abstract from the renegotiation possibilities that are provided under GATT/WTO rules. But Bagwell and Staiger (1999, 2002) argue that the particular renegotiation provisions included in GATT Article XXVIII have the effect of ensuring that no country can be forced in a bilateral GATT/WTO bargain to accept greater trade volume than it desires at the given terms of trade/world price. This is an interesting feature of the GATT/WTO bargaining setting to consider in this context, because intuitively the possibility of Article XXVIII renegotiation could diminish the amount of negotiated trade liberalization that bargaining partners can achieve. In light of our results above concerning bilateral tariff bargaining outcomes in the absence of MFN and in its presence, this suggests that the possibility of Article XXVIII renegotiation might diminish the amount of excessive liberalization

in the absence of MFN and thereby move the outcome toward the efficiency frontier, while it could either move the outcome toward or away from the efficiency frontier in the presence of MFN.[32]

While the intuition for the above statements seems strong, the method of analysis to confirm this intuition would have to be quite different from that which we have pursued above. In the preceding sections, we rely heavily on the fact that at a bilateral tariff bargaining solution each bilateral satisfies a tangency condition for the indifference curves of the two bargaining parties when the policies determined in the other bilateral are taken as fixed. With Article XXVIII renegotiation, however, this tangency condition will in general not be met. A different approach to constructing perturbations from the tariff bargaining solution is therefore necessary to proceed with an analysis that includes the possibility of Article XXVIII renegotiation. We leave this important task for future research.

10 Conclusion

We consider a three-country, two-good model of bilateral tariff negotiations where each country is affected by the outcomes achieved in each bilateral negotiation. We focus on weak-rule settings characterized by either no rules or only the MFN rule. Discriminatory tariffs are allowed in the no-rules case, whereas tariffs must be non-discriminatory in the MFN-only case. For the case with no rules, we characterize the negotiated tariffs that are predicted by the Horn-Wolinsky (1988) solution. When the MFN rule is imposed, we characterize the MFN tariffs that are predicted by a Nash bargaining solution between the home country and its principal supplier when the tariff policy of the other foreign country is taken as exogenous.

In both cases, our objective is to characterize the efficiency properties of the negotiated tariffs. For the no-rules case, we show that starting from an interior Horn-Wolinsky solution we can construct a Pareto-improving perturbation under which all tariffs are increased. We also characterize the necessary features of Pareto-improving perturbations, showing that, if at least one country strictly gains under such a perturbation, then there must exist a bilateral relationship in which at least one tariff rises while the other tariff in that relationship can fall but only to a limited extent. Based on these findings, we conclude for the no-rules case with discriminatory tariffs that simultaneous bilateral tariff negotiations are associated with excessive liberalization when judged relative to the preferences of all countries. By contrast, for the MFN-only case, we show for one important case that, starting from an MFN solution, we can construct a Pareto-improving perturbation under which all tariffs are reduced. In this way, our findings thus indicate that the MFN rule provides a partial counterbalance to the forces that result in inefficiently low tariffs when discriminatory tariffs are permitted. We also provide characterizations of the efficiency of MFN solutions relative to the MFN-constrained efficiency frontier. Finally, we provide numerical characterizations that concretely illustrate and further develop our findings for an endowment economy with Cobb-Douglas preferences.

---

[32]Renegotiation under GATT Article XXVIII is conducted with respect to MFN tariff bindings, but it is also interesting to consider the implications of bilateral renegotiation in a setting with discriminatory tariffs.

Our work also contributes at a methodological level. The Nash-in-Nash approach of the Horn-Wolinsky solution underlies a large and important body of applied work in Industrial Organization that studies surplus division in bilateral oligopoly settings. Our work here provides a theoretical foundation for related studies in International Trade that address bilateral tariff negotiations. In addition to such applied work, our work motivates further examination of the micro-foundation of the Nash-in-Nash solution for settings in which negotiated outcomes go beyond surplus division and impact efficiency.

11 Appendix

To establish conditions for the existence of a Horn-Wolinsky solution as defined in the text, we define a generalized game with infinite strategy spaces and two players. The objective of player i, where $i = 1, 2$, is to select τ^i and τ^{*i} so as to maximize the Nash bargaining solution objective for the bargaining relationship between the home country and foreign country $*i$. Each player i, however, must also select τ^i and τ^{*i} from the space of feasible tariffs that satisfy participation constraints, as captured by the weak inequalities stated in the text. We note that the participation constraints for player i are affected by the strategy choices of player j, where $j = 1, 2$ and $i \neq j$.

Formally, player $i = 1, 2$ has a strategy $s^i \equiv (\tau^i, \tau^{*i})$, where $s^i \in S \equiv [\underline{\tau}, \overline{\tau}]^2$ with $(\underline{\tau}, \overline{\tau}) \in \Re^2$ and $0 < \underline{\tau} < \overline{\tau}$. Player i has the payoff function $g^i(s^1, s^2)$, where

$$
\begin{aligned}
g^1(s^1, s^2) &\equiv \Delta W^1(s^1, s^2; s_0^1) \cdot \Delta W^{*1}(s^1, s^2; s_0^1) \\
g^2(s^1, s^2) &\equiv \Delta W^2(s^1, s^2; s_0^2) \cdot \Delta W^{*2}(s^1, s^2; s_0^2),
\end{aligned}
$$

and where $s_0^i \equiv (\tau_0^i, \tau_0^{*i}) \in S$, $i = 1, 2$, are exogenously given. Finally, in recognition of the participation constraints, we further restrict s^i to a subset $\gamma^i(s^j)$ of S, where we define

$$
\begin{aligned}
\gamma^1(s^2) &\equiv \{s^1 \in S \mid W(s^1, s^2) \geq W(s_0^1, s^2) \text{ and } W^{*1}(s^1, s^2) \geq W^{*1}(s_0^1, s^2)\} \\
\gamma^2(s^1) &\equiv \{s^2 \in S \mid W(s^1, s^2) \geq W(s^1, s_0^2) \text{ and } W^{*2}(s^1, s^2) \geq W^{*2}(s^1, s_0^2)\}.
\end{aligned}
$$

We now say that a pair $(\hat{s}^1, \hat{s}^2)$ is a generalized Nash equilibrium if, for all $i, j = 1, 2$ with $i \neq j$, $\hat{s}^i \in \gamma^i(\hat{s}^j)$; $g^1(\hat{s}^1, \hat{s}^2) \geq g^1(s^1, \hat{s}^2)$ for all $s^1 \in \gamma^1(\hat{s}^2)$; and $g^2(\hat{s}^1, \hat{s}^2) \geq g^2(\hat{s}^1, s^2)$ for all $s^2 \in \gamma^2(\hat{s}^1)$. The Horn-Wolinsky solution may now be understood as a generalized Nash equilibrium for the two-person generalized game defined here.

We recall that $S \equiv [\underline{\tau}, \overline{\tau}]^2$ and that $W(\boldsymbol{\tau})$ and $W^{*i}(\boldsymbol{\tau})$ are continuously differentiable for $i = 1, 2$. It follows that S is a nonempty, compact and convex subset of Euclidian space and that, for $i = 1, 2$, $g^i(s^1, s^2)$ is continuous in (s^1, s^2). According to Debreu's (1952, 1983) theorem, a pure strategy generalized Nash equilibrium exists for the generalized two-person game defined here if for $i, j = 1, 2$ and $i \neq j$, (a) $g^i(s^1, s^2)$ is quasiconcave in s^i, and (b) $\gamma^i(s^j)$ is upper and lower hemicontinuous, convex valued and nonempty valued. Equivalently, conditions (a) and (b) ensure the existence of a Horn-Wolinsky solution for the model defined in the text.

Note that condition (a) imposes quasiconcavity in (τ^i, τ^{*i}) for the Nash Bargaining solution objective, $\Delta W^i(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}; \tau_0^i, \tau_0^{*i}) \cdot \Delta W^{*i}(\tau^1, \tau^{*1}, \tau^2, \tau^{*2}; \tau_0^i, \tau_0^{*i})$, rather than for the individual welfare functions.[33] We can show that condition (b) holds if, for each $i, j = 1, 2$ with $i \neq j$ and for any $s^j \in S$, $W(\boldsymbol{\tau})$ and $W^{*i}(\boldsymbol{\tau})$ are strictly quasiconcave in s^i. It is direct to verify that $\gamma^i(s^j)$ is nonempty (since s_0^i is a member), convex valued (since $W(\boldsymbol{\tau})$ and $W^{*i}(\boldsymbol{\tau})$ are quasiconcave in s^i)

---

[33]In the supplementary materials provided in our online appendix, we consider a particular representation of the model as an endowment economy with Cobb-Douglas preferences. Under the assumption that each government maximizes the indirect utility of a representative agent in its country, we use a numerical example to illustrate that condition (a) plausibly holds for examples of interest.

and upper hemicontinuous (since $W(\boldsymbol{\tau})$ and $W^{*i}(\boldsymbol{\tau})$ are continuous). Using the strict quasiconcavity of $W(\boldsymbol{\tau})$ and $W^{*i}(\boldsymbol{\tau})$ in s^i, we can also show that $\gamma^i(s^j)$ is lower hemicontinuous in our setting. Finally, we also note that the conditions stated here ensure existence but do not ensure interiority.

12 References

Bagwell, Kyle and Robert W. Staiger (1999), "An Economic Theory of GATT," *American Economic Review* 89(1): 215-48.

Bagwell, Kyle and Robert W. Staiger (2002), *The Economics of the World Trading System*, Cambridge, MA: The MIT Press.

Bagwell, Kyle and Robert W. Staiger (2005), "Multilateral Trade Negotiations, Bilateral Opportunism and the Rules of GATT/WTO," *Journal of International Economics* 67(2): 268-94.

Bagwell, Kyle and Robert W. Staiger (2010), "Backward Stealing and Forward Manipulation in the WTO," *Journal of International Economics* 82(1): 49-62.

Bagwell, Kyle and Robert W. Staiger (2016), "Multilateral Trade Bargaining and Dominant Strategies," NBER Working Paper 22842.

Bagwell, Kyle, Robert W. Staiger and Ali Yurukoglu (2017), "Multilateral Trade Bargaining: A First Look at the GATT Bargaining Records," working paper.

Beckett, Grace (1941), *The Reciprocal Trade Agreements Program*. New York: Columbia University Press.

Chan, Jackie M. L. (2015), "Strategic Tariff Bargaining in the GATT and Gradualism," working paper.

Chipty, Tasneem and Chris M. Snyder (1999), "The Role of Firm Size in Bilateral Bargaining: A Study of the Cable Television Industry," *Review of Economics and Statistics*, 81(2), 326-40.

Collard-Wexler, Allan, Gautam Gowrisankaran, and Robin S. Lee (2016), "'Nash-in-Nash Bargaining: A Microfoundation for Applied Work," working paper.

Crawford, Greg, Robin S. Lee, Michael D. Whinston and Ali Yurukoglu (2015), "The Welfare Effects of Vertical Integration in Multichannel Television Markets," working paper.

Crawford, Greg and Ali Yurukoglu (2012), "The Welfare Effects of Bundling in Multi-channel Television Markets," *American Economic Review*, 102(2), 643-85.

Cremer, Jacques and Michael H. Riordan (1987), "On Governing Multilateral Transactions with Bilateral Contracts," *RAND Journal of Economics*, 18, 436-51.

Dasgupta, Partha Sarathi and Eric S. Maskin (2015), "Debreu's Social Equilibrium Existence Theorem," *Proceedings of the National Academy of Sciences*, 112(52), 15769-70.

Debreu, Gerard (1952), "A Social Equilibrium Existence Theorem," *Proceedings of the National Academy of Sciences*, 38(10), 886-93.

Debreu, Gerard (1983), "Existence of Competitive Equilibrium," in *Handbook of Mathematical Economics*, vol. 2, Kenneth Arrow and Michael Intrilligator (eds.), New York: North Holland.

Gowrisankaran, Gautum, Aviv Nevo and Robert Town (2015), "Mergers when Prices Are Negotiated: Evidence from the Hospital Industry," *American Economic Review*, 105(1), 172-203.

Grennan, Matthew (2013), "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices," *American Economic Review*, 103(1), 147-77.

Hart, Oliver and Jean Tirole (1990), "Vertical Integration and Market Foreclosure," *Brookings Papers on Economic Activity, Microeconomics*, 205-76.

Ho, Kate and Robin S. Lee (2017), "Insurer Competition in Health Care Markets," *Econometrica*, 85(2): 379-417.

Horn, Henrik and Asher Wolinsky (1988), "Bilateral Monopolies and Incentives for Merger," *RAND Journal of Economics* 19(3): 408-419.

McAfee, R. Preston and Marius Schwartz (1994), "Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity," *American Economic Review*, 84, 210-30.

McAfee, R. Preston and Marius Schwartz (1995), "The Non-existence of Pairwise-proof Equilibria," *Economics Letters*, 49, 251-59.

Ossa, Ralph (2014), "Trade Wars and Trade Talks with Data," *American Economic Review*, 104(12), 4104-46.

Pomfret, Richard (1997), *The Economics of Regional Trading Arrangements*, New York, NY: Oxford University Press.

Rohlfing, Ingo (2009), "Discrimination in International Trade: A Different Perspective," *Acta Politica*, 44, 192-210.

Saggi, Kamal and Halis Murat Yildiz (2010), "Bilateralism, Multilateralism and the Quest for Global Free Trade," *Journal of International Economics*, 81(1), 26-37.

Saggi, Kamal, Halis Murat Yildiz and Alan Woodland (2013), "On the Relationship between Preferential and Multilateral Trade Liberalization: The Case of Customs Unions," *American Economic Journal: Microeconomics*, 5(1), 63-99.

Tasca, Henry. J. (1938), *The Reciprocal Trade Policy of the United States: A Study in Trade Philosophy*, Philadelphia, PA: University of Pennsylvania Press.

Tavares, Samia C. (2006), "The Political Economy of the European Customs Classification," *Public Choice*, 129(1-2), 107-30.
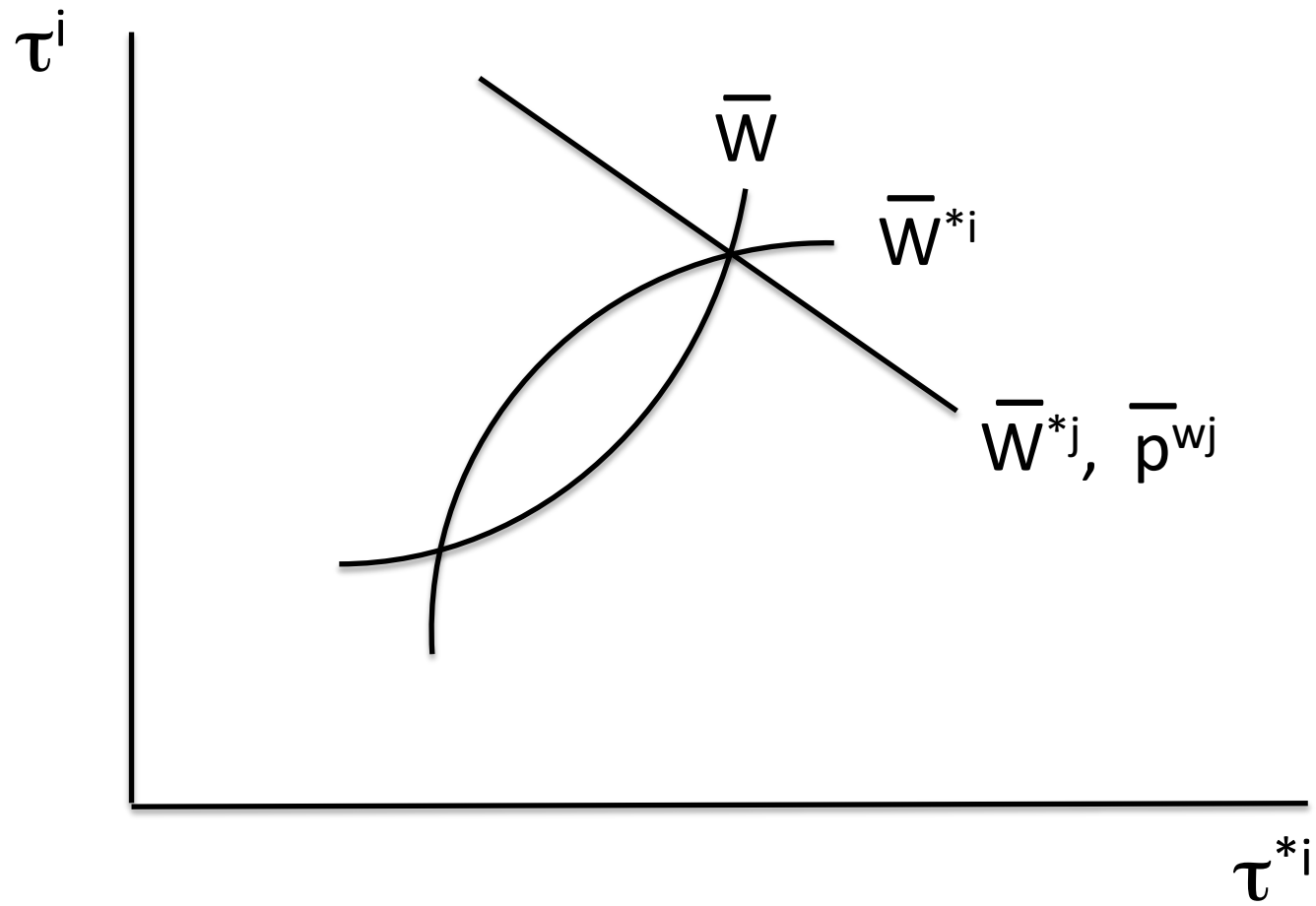
Figure 1
Efficient Tariffs

Figure 2
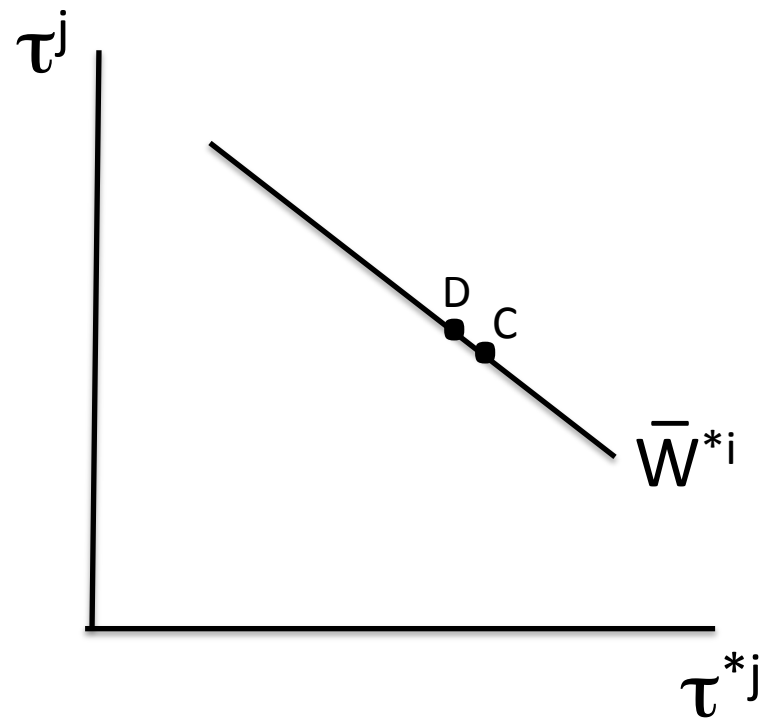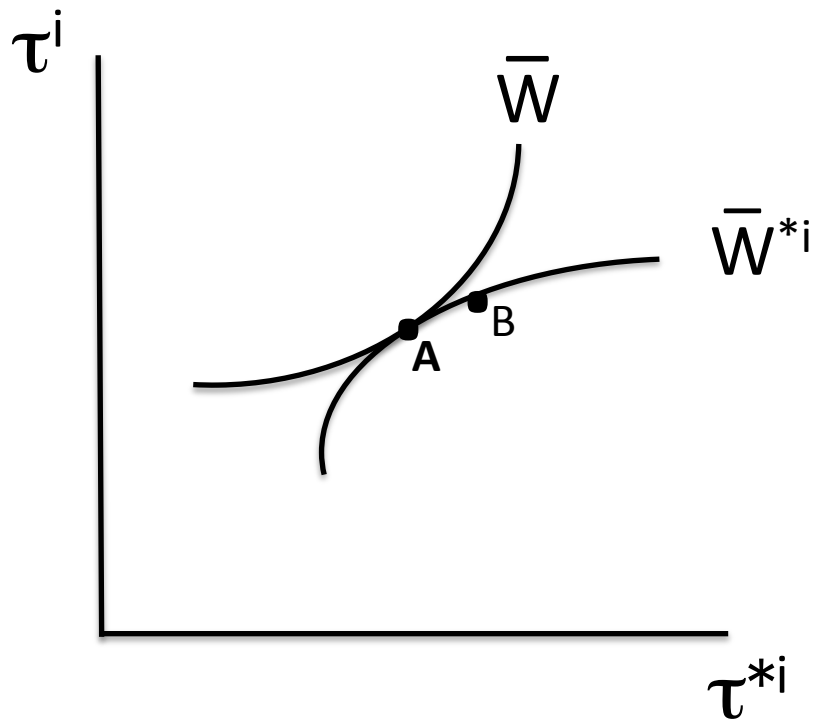**Two-step Perturbation**
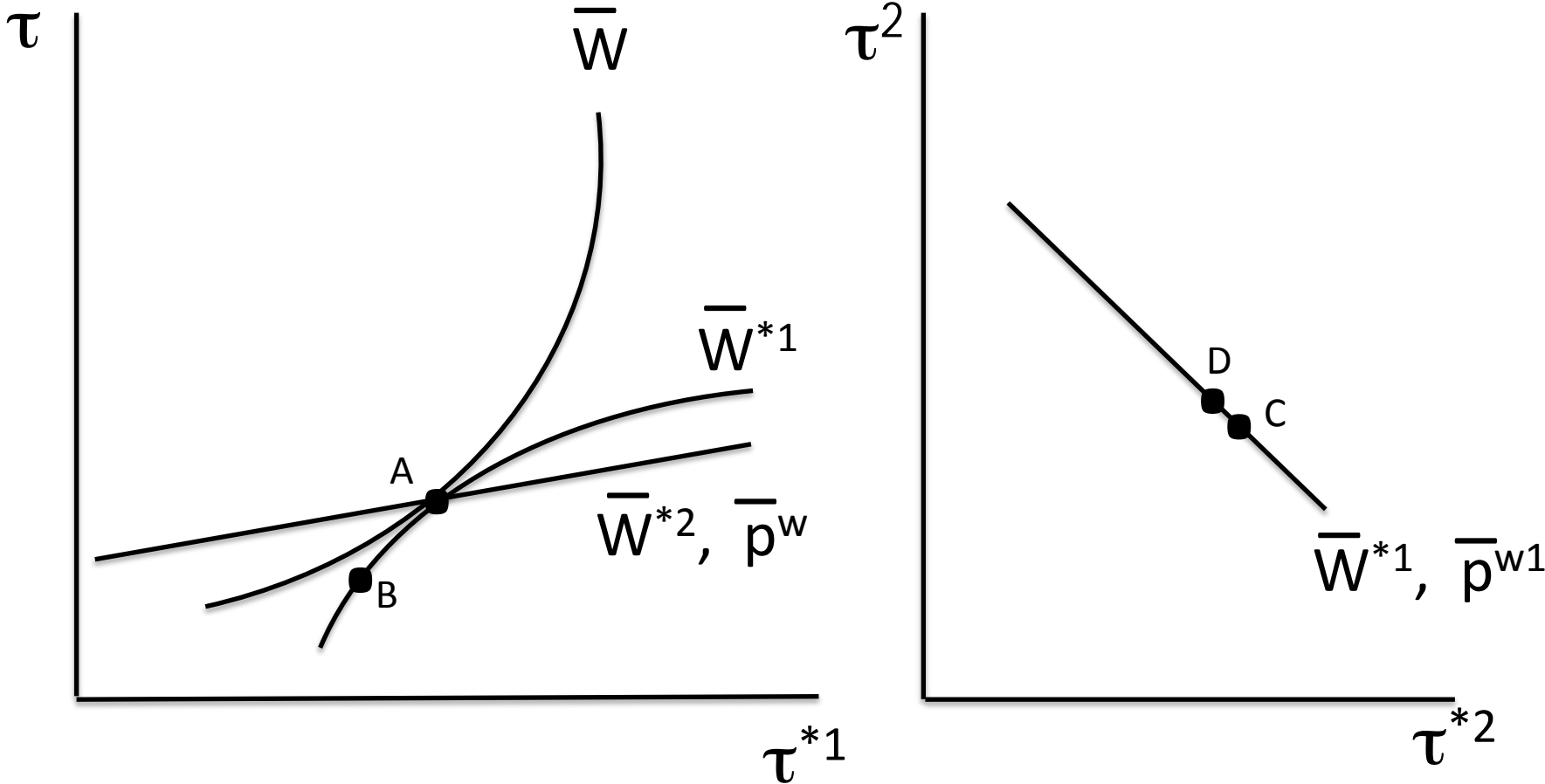
Figure 3
Case 1: MFN tariffs too high
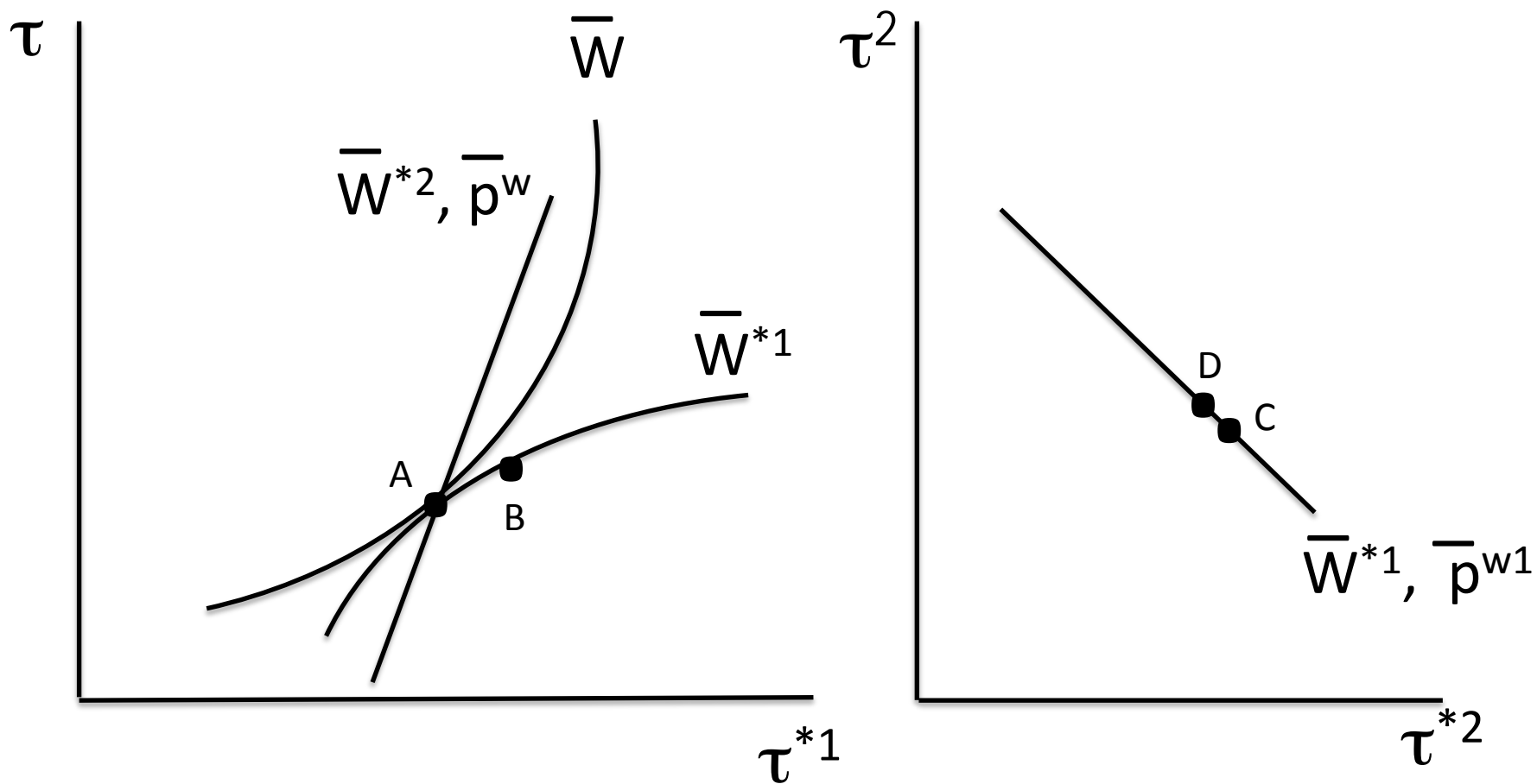
Figure 4
Case 2: MFN tariffs too low

Figure 5
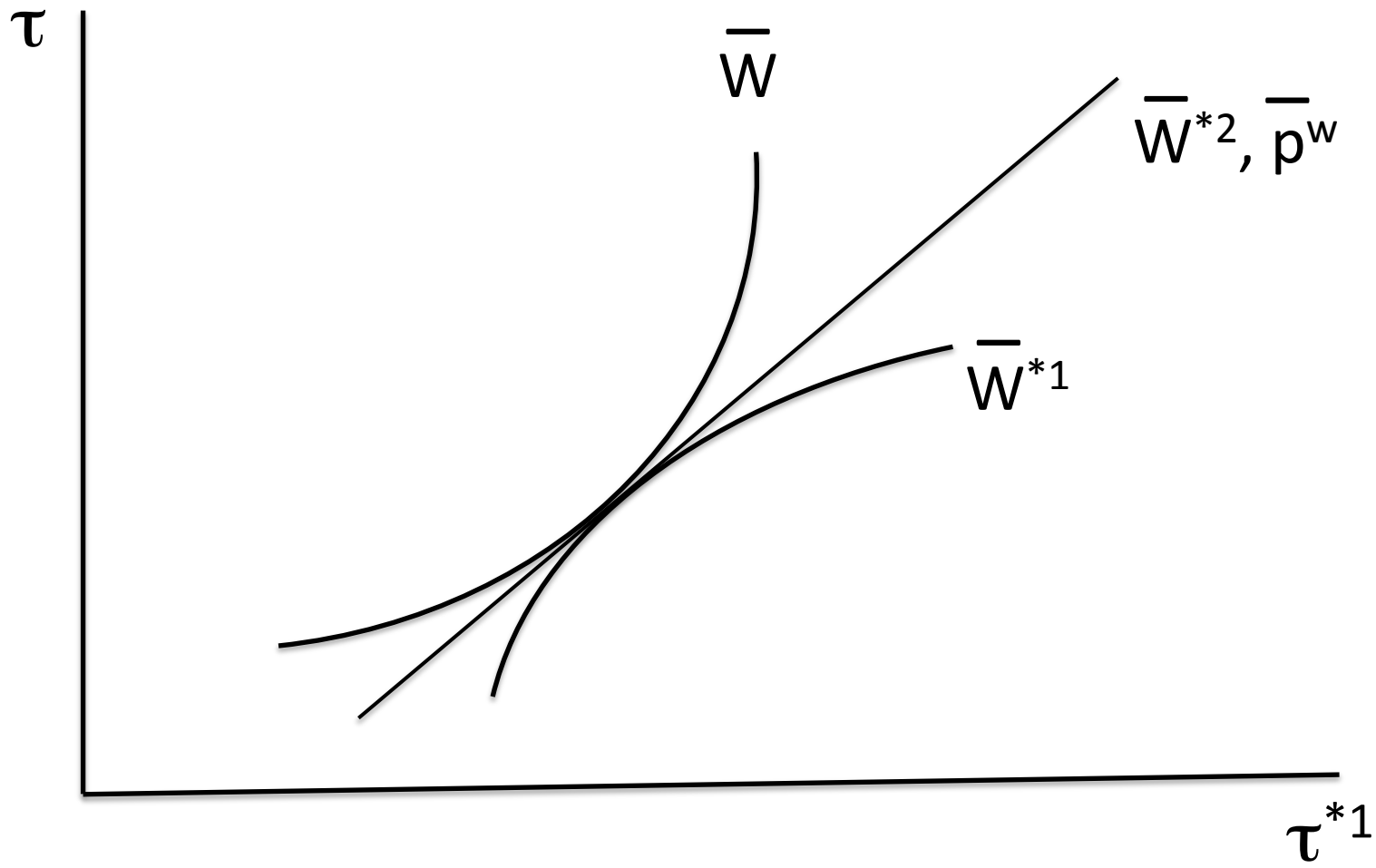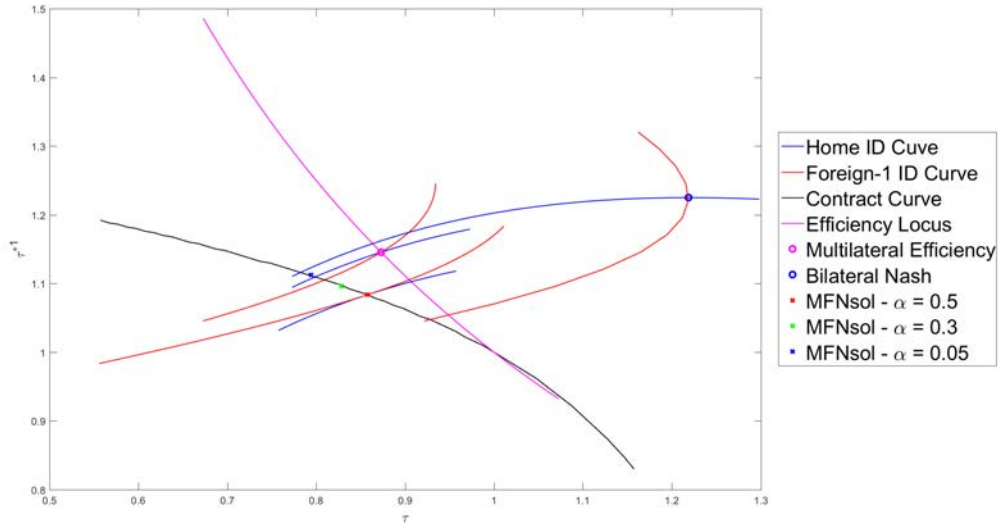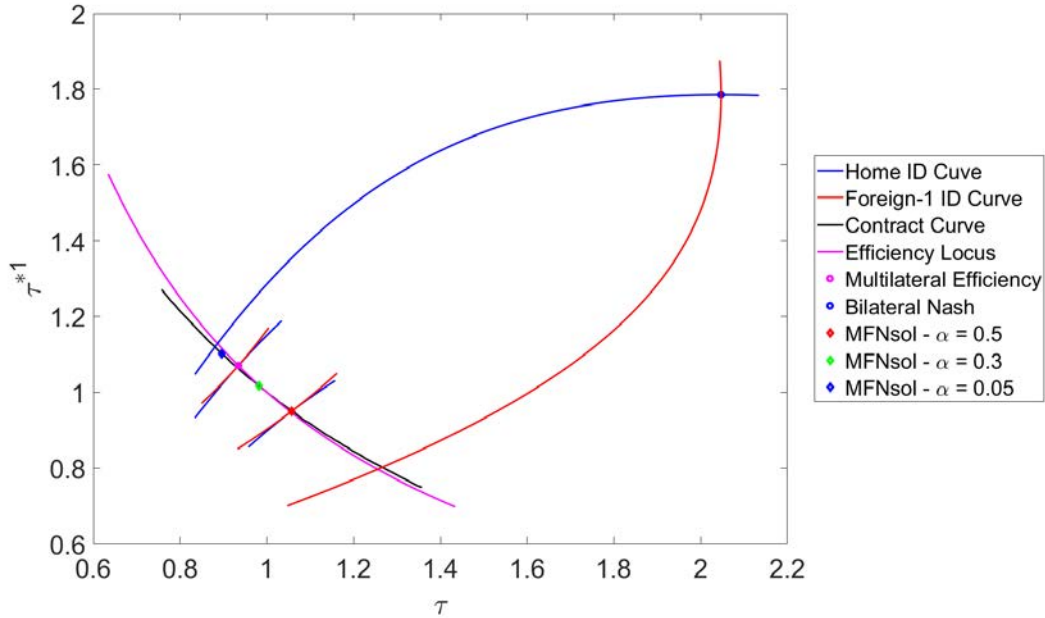## Case 3: MFN tariffs at preferred volumes

Figure 6: Illustration of the MFN Solution

Notes: The figure illustrates the case for $\beta = 0.6, \gamma = 0.1, \lambda = 0.03$. The red and blue curves are indifference curves for home and foreign country *1 respectively. The black curve is the bargaining locus, and the contract curve is the portion of the black curve that falls between the indifference curves. The pink curve is the multilateral efficiency locus defined as $\tau \cdot \tau^{*1} = 1$. The pink dot is the multilaterally efficient point and the blue dot corresponds to bilateral Nash Equilibrium. Three dots along the contract curve correspond to MFN-solutions for different bargaining weights.

Figure 7: Illustration of the MFN Solution

Notes: The figure illustrates the case for $\beta = 0.9, \gamma = 0.2, \lambda = 0.9$. The red and blue curves are indifference curves for home and foreign country *1 respectively. The black curve is the bargaining locus, and the contract curve is the portion of the black curve that falls between the indifference curves. The pink curve is the multilateral efficiency locus defined as $\tau \cdot \tau^{*1} = 1$. The pink dot is the multilaterally efficient point and the blue dot corresponds to bilateral Nash Equilibrium. Three dots along the contract curve correspond to MFN-solutions for different bargaining weights.