Title
Essays on international trade agreements and contracts under renegotiation

Author
Buzard, Kristina L.

Publication Date
2012

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays on International Trade Agreements and Contracts under Renegotiation

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Economics

by

Kristina L. Buzard

Committee in charge:

Professor Joel Watson, Co-Chair
Professor James Rauch, Co-Chair
Professor J. Lawrence Broz
Professor Peter Cowhey
Professor Marc-Andreas Muendler

2012

The dissertation of Kristina L. Buzard is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Co-Chair

University of California, San Diego

2012

DEDICATION

To David and my parents, whose love and support mean the world.

EPIGRAPH

It's better to be roughly right than precisely wrong.

—John Maynard Keynes

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Joel Watson for his unquenchable enthusiasm and unconditional support. His guidance and confidence have enabled me to achieve so many dreams.

I have benefited enormously from the wisdom, knowledge and generosity of an uncommonly wonderful dissertation committee and so many other superb faculty. To have been surrounded by such *kindness*: this is the debt I must perhaps work the hardest to repay, and yet I know that I will never fully be able to balance the ledger.

I would also like to acknowledge all the graduate students who have shared so many steps of this journey. I have learned so much with you and laughed so much with you and I cannot imagine doing this with anyone else. I don't know how it would be possible to do this without such an amazing community, and I feel so blessed never to have to find out. To the Tuesday noon kids—you seem to have accomplished maybe the toughest feat—helping me to stay sane.

To all those from my time at Drexel, the College of Europe, and the Federal Reserve Bank of Philadelphia who helped to prepare me for this stage of the journey and who have continued to love and support me throughout, I will be forever grateful. It is so clear that I never could have tackled the work represented here without you. In particular, I do not see how any of this would have been possible without Bob Hunt, Jerry Carlino, Dr. De Souza, Dr. P. or Dr. Thweatt, may she rest in peace.

Finally, to the members of the kewlest SCC ever: you made me one of you, and you made this home. No matter what life, and school, tossed at me, I could come to you each week, and you reminded me why it all *mattered*. From the very deepest place within me, thank you.

Chapter 1, in full, is a reprint of the material as it appears in *Theoretical Economics*, 7, 2012, Buzard, Kristy; Watson, Joel, 2012. The dissertation author was a principal author of this paper.

VITA

2001	Bachelor of Arts, International Area Studies *summa cum laude*, Drexel University, Philadelphia, PA
2002	Master of European Economic Studies, College of Europe, Bruges, Belgium
2002-2007	Economic Analyst, Research Department, Federal Reserve Bank of Philadelphia, Philadelphia, PA
2007-2012	Graduate Teaching Assistant, University of California, San Diego
2012	Doctor of Philosophy, Economics, University of California, San Diego

PUBLICATIONS

"Contract, Renegotiation, and Hold Up: Results on the Technology of Trade and Investment," *Theoretical Economics*, 7, 2012 (with Joel Watson).

"Spatial Distribution of R&D Activity in the United States," *Handbook of Economic Geography and Industry Studies*, F. Giarratani, G. Hewings, and P. McCann, eds. Cheltenham: Edward Elgar Publishing, forthcoming. (with Gerald A. Carlino).

"The Geography of Research and Development Activity in the U.S.," *Business Review*, Federal Reserve Bank of Philadelphia, Third Quarter 2008 (with Gerald A. Carlino).

ABSTRACT OF THE DISSERTATION

Essays on International Trade Agreements
and Contracts under Renegotiation

by

Kristina L. Buzard

Doctor of Philosophy in Economics

University of California, San Diego, 2012

Professor Joel Watson, Co-Chair
Professor James Rauch, Co-Chair

The first chapter of the dissertation addresses general issues in contracting with external enforcement. We study a contracting environment with specific investments in which renegotiation, and therefore hold-up, is possible. We show that taking account of the precise nature of trading and investment technologies is important for accurately determining the trading relationships in which efficient investment and trade will occur and that careful modeling of institutional detail and the information available to private parties and the external enforcement body (e.g. a court) are key.

The second chapter presents a model of international trade agreements in which

domestic policy-making power is shared between executive and legislative branches of government. Acknowledging the complexity of the legislative process as well as its susceptibility to lobbying reveals a political commitment role for trade agreements in that executives can use them to reduce incentives for lobbying so that the legislatures can better withstand political pressure. This helps explain the result from tests of the Grossman and Helpman (1994) model that there is too much protection relative to contributions given estimates of governments' social-welfare weights: I predict that contribution levels may in fact be low *because* tariffs have been raised to prevent political pressure and the increased risk of a trade disruption it engenders.

The third chapter extends this model to a repeated-game framework, replacing the assumption of external enforcement with self-enforcing promises of future cooperation. Here, the inability of actors to make commitments affects the design of trade agreements in two ways: executives must not only take into account the legislatures' lobbying-driven propensity to revoke delegation and break the agreement, but also be robust to the executives' own incentives to renegotiate out of any punishment scheme. The design of the dispute resolution mechanism that makes the optimal punishment incentive compatible must balance two, often-conflicting, objectives: longer punishment periods help to enforce cooperation by increasing the costs of defecting from the agreement, but because the lobbies prefer the punishment outcome, this also incentivizes lobbying effort and with it the political pressure to break the agreement. Thus the model generates new predictions for the optimal design of mechanisms for resolving the disputes that arise in the course of trade-agreement relationships.

Chapter 1

Contract, renegotiation, and holdup: Results on the technology of trade and investment

Theoretical Economics 7 (2012), 283–322 1555-7561/20120283

Contract, renegotiation, and holdup: Results on the technology of trade and investment

Kristy Buzard
Department of Economics, University of California, San Diego

Joel Watson
Department of Economics, University of California, San Diego

This paper examines a class of contractual relationships with specific investment, a nondurable trading opportunity, and renegotiation. Trade actions are modeled as individual and trade-action-based option contracts ("nonforcing contracts") are explored. The paper introduces the distinction between *divided* and *unified* investment and trade actions, and it shows the key role this distinction plays in determining whether efficient investment and trade can be achieved. Under a nonforcing *dual-option contract*, the party without the trade action is made the residual claimant with regard to the investment action, which induces efficient investment in the divided case. The unified case is more problematic: here, efficiency is typically not attainable, but the dual-option contract is still optimal in a wide class of settings. More generally, the paper shows that, with ex post renegotiation, constraining parties to use "forcing contracts" implies a strict reduction in the set of implementable value functions.

Keywords. Contract, renegotiation, holdup, forcing contracts, nonforcing contracts, specific investment, technology of trade, mechanism design.

JEL classification. C70, D23.

The holdup problem arises in situations in which contracting parties can renegotiate their contract between the time they make unverifiable relation-specific investments and the time at which they can trade.[1] The severity of the holdup problem depends critically on the productive technology and on the timing of renegotiation opportunities. This paper contributes to the literature by examining how the nature of the "trade

Kristy Buzard: kbuzard@ucsd.edu
Joel Watson: jwatson@ucsd.edu
The authors thank the following people for their insightful comments: the anonymous referees, Nageeb Ali, Jeff Ely, Bob Evans, David Miller, Ben Polak, Larry Samuelson, Joel Sobel, and seminar participants at Columbia, Florida International, UCSD, USC, SWET, and Yale. Part of the analysis reported here was completed while Watson was a visitor at the Cowles Foundation, Yale, and he is grateful for the support from Cowles.

[1] Che and Sákovics (2008) provide a short overview of the holdup problem, which was first described by Klein et al. (1978), and Williamson (1975, 1979). Analysis was provided by Grout (1984), Grossman and Hart (1986), and Hart and Moore (1988).

DOI: 10.3982/TE818

action" in a contractual relationship influences the prospects for achieving an efficient outcome. We introduce a new distinction—whether the party who invests also is the one who consummates trade—that plays an important role in determining the outcome of the contractual relationship.

So that we can describe our modeling exercise more precisely, consider an example in which contracting parties Al and Zoe interact as follows. First Al and Zoe meet and write a contract that has an externally enforced element. Then one of them makes a private investment choice, which influences the *state* of the relationship. The state is commonly observed by the contracting parties, but is not verifiable to the external enforcer. Al and Zoe then send individual public messages to the external enforcer. After this, they have an opportunity to renegotiate their contract: this is called ex post renegotiation because it occurs after messages. Finally, the parties have a one-shot opportunity to trade and they also obtain external enforcement. Trade is verifiable to the external enforcer.

Because the investment is unverifiable, the investor cannot be directly rewarded for choosing the efficient investment level. Instead, incentives hinge on how the terms of trade can be made sensitive to the investment choice. Typically a conflict arises between the parties' joint interests prior to investment and their joint interests following investment and messages. In particular, investment incentives may be strengthened by specifying an inefficient trade action ex post in some off-equilibrium-path contingencies. But parties then would have the joint incentive to renegotiate and divide the surplus according to their bargaining power (holdup). Because parties rationally anticipate the renegotiation, the incentives to invest are distorted.

The description above obviously leaves the mechanics of trade and enforcement ambiguous. In reality, the parties have individual actions that determine whether and how trade is consummated. Let us suppose that Al selects the individual trade action, which we call a. This could be a choice of whether to deliver or to install an intermediate good, for example. We then have an *individual-action model*, whereby Al chooses a and the external enforcer compels a transfer t as a function of a and the messages that the parties sent earlier. In contrast, a *public-action model* (or *external-action model*) combines the trade action and the monetary transfer into a single public action (a, t) that is assumed to be taken by the external enforcer. With this modeling approach, the contract specifies how the public action is conditioned on the parties' messages.

Although the public-action model may typically be a bit unrealistic, it is simple and lends itself to elegant mechanism-design analysis (for example, as in Maskin and Moore 1999 and Segal and Whinston 2002). Alternatively, Watson (2007) demonstrates that analysis of the individual-action model can be straightforward as well. He also shows that the public-action model is equivalent to examining individual trade actions, but constraining attention to "forcing contracts" in which the external enforcer induces a particular trade action as a function of messages sent by the parties (so the trade action is constant in the state). Watson (2007) provides an example in which the restriction to forcing contracts has strictly negative efficiency consequences.

We deepen the examination of nonforcing contracts by investigating their efficacy in the context of different technologies of trade and investment. Specifically, we introduce the distinction between *divided* and *unified* investment and trade actions. In the divided case, the investment and trade actions are chosen by different parties (Al takes the trade action and Zoe makes the investment). In the unified case, the investment and trade actions are selected by the same party (Al does both). We show that the prospects for inducing efficient investment and trade are very different in the divided and unified cases. In fact, the efficient outcome can always be achieved in the divided case (assuming investment has no immediate benefits), but typically cannot be achieved in the unified case.

Our analysis also highlights a simple contractual form that we call a *dual-option contract*. With the dual-option, Zoe sends a message that can be interpreted as a requested trade action or declaration of the state, and Al's subsequent trade action also serves as an option. We show that a dual-option contract is optimal in a large class of contractual relationships. For instance, it can be used to make Al's payoff constant in the state, gross of any investment costs, so that Zoe becomes the residual claimant with respect to the investment choice. This implies the efficiency result for the divided case. The dual option is also useful in the unified case, even though the efficient outcome typically cannot be achieved; specifically, we show that in a class of settings with a deterministic state, the dual-option contract is optimal.

Our analysis utilizes mechanism-design techniques. With both the individual-action and public-action modeling approaches, analysis of the contractual problem centers on calculating the set of implementable value functions from just after the state is realized (before messages are sent). Formally, an implementable value function is the state-contingent continuation value that results in equilibrium for a given contract. We provide simple tools to calculate the "punishment values" that determine the implementable sets for the class of relationships we analyze here. We use these tools to characterize optimal contracts and to find bounds on the set of implementable value functions.

In addition to the results on the divided and unified cases and dual-option contracts, we provide a general result on the comparison of forcing and nonforcing contracts, which shows that Watson's (2007) conclusions are robust over a large class of contractual relationships. In particular, in the important setting of ex post renegotiation described above, limiting attention to forcing contracts reduces the set of state-contingent continuation values. This does not mean that a more efficient outcome can always be achieved when actions are modeled as individual (because efficiency depends on what region of the implementable value set is relevant for giving appropriate investment incentives), but it underscores the importance of modeling trade actions as individual.

This is particularly salient for the setting of *cross/cooperative investment* (Che and Hausch 1999), where the investment by one party increases the benefit to the other party of subsequent trade. The literature has regarded cross-investment settings as especially prone to the holdup problem. Che and Hausch (1999) show that the optimal forcing contract is often "null" and leads to underinvestment. Our results establish that nonforcing contracts offer a significant improvement in efficiency, and our distinction between unified and divided investment and trade actions gives a basis for deeper analysis.

In the class of trade technologies that we study here, a single player (player 1, Al above) has the trade action. Examples of real settings with this property are contractual relationships in which the seller provides a service or good that does not require the buyer's involvement (such as consulting, advertising, and some types of construction). In these settings, the seller has the trade action. Other settings with unilateral trade actions are ones in which the seller is the investor, production is inherent in the seller's investment, and trade is determined by whether the buyer installs or otherwise adopts the intermediate good; an example is specialized software. In these settings, the buyer has the trade action. We discuss the extension to multilateral trade actions in the Conclusion (Section 6).

The only assumption required for our first result, on making player 2 (Zoe) the residual claimant, is that investment does not confer a direct gain for some minimal trade action (an assumption satisfied by the most prominent models in the holdup literature). The key economic assumption behind our other results is that player 1's utility is supermodular as a function of the state and the trade action. That is, this player's marginal value of the trade action is monotone in the state. Our other assumptions are mainly weak technical conditions that guarantee well defined maxima, nontrivial settings, and the like. We argue that these conditions are likely to hold in a wide range of applications and that they are consistent with typical assumptions in the literature. Our result about the optimality of the dual-option contract in the unified case requires some additional assumptions on the technology of investment and trade.

The rest of the paper proceeds as follows. In the next section, we provide the details of the model. Section 2 presents an example that illustrates our main results. Section 3 contains our general results on optimal contracts and outcomes in the divided and unified cases. Readers interested in getting all of the basic ideas without the technical details can proceed from Section 3 straight to the Conclusion. Section 4 provides an overview of the basic tools for general analysis, which mostly restates material in Watson (2007). Section 5 contains our result on the difference in implementable sets based on variations regarding when renegotiation can occur and whether one restricts attention to forcing contracts. The Conclusion contains more discussion about the holdup problem and cross investment, as well as notes on the case of durable trading opportunities and multilateral trade actions. Most of the technical material and all of the proofs are contained in the Appendices.

1. The theoretical framework

We look at the same class of contracting problems and use the same notation as in Watson (2007), except that we add a bit of structure on the trade technology to focus our analysis. In particular, we examine the case in which a single player has a trade action. Throughout the paper, we use the convention of labeling the player with the trade action as player 1, and we call the other player 2. These two players are the parties engaged in a contractual relationship with a nondurable trading opportunity and external enforcement. Their relationship has the following payoff-relevant components, occurring in the order shown.

The *state of the relationship* θ. The state represents unverifiable events that are assumed to happen early in the relationship. The state may be determined by individual investment decisions and/or by random occurrences, depending on the and selecting the trade actions at date 6. At date 8, the external enforcer compels transfers. At odd-numbered dates, the players make joint contracting decisions—establishing a contract at date 1 and possibly renegotiating it later.

The contract has an externally enforced component consisting of (i) feasible message spaces M_1 and M_2 and (ii) a *transfer function* $y : M \times A \to \mathbb{R}^2$ that specifies the transfer t as a function of the verifiable items m and a. That is, having seen m and a, the external enforcer compels transfer $t = y(m, a)$. The contract also has a self-enforced component, which specifies how the players coordinate their behavior for the times at which they take individual actions. Renegotiation of the contract amounts to replacing the original transfer function y with some new function y', in which case y' is the one submitted to the external enforcer at date 8.

We initially assume—and maintain throughout Sections 2 and 3—that the players can freely renegotiate at dates 3, 5, and 7. Renegotiation at date 5 is called *ex post renegotiation*. At date 3 it is called *interim renegotiation*.[3]

The players' individual actions at dates 2, 4, and 6 are assumed to be consistent with sequential rationality; that is, each player maximizes his expected payoff, conditional on what occurred earlier and on what the other player does, and anticipating rational behavior in the future. The joint decisions (initial contracting and renegotiation at odd-numbered periods) are assumed to be consistent with a cooperative bargaining solution in which the players divide surplus according to fixed bargaining weights π_1 and π_2 for players 1 and 2, respectively. The bargaining weights are nonnegative, sum to 1, and are written $\pi = (\pi_1, \pi_2)$. The negotiation surplus is the difference between $\gamma(\theta)$ and the joint value that would result if the players fail to reach an agreement, where the disagreement point is given by an equilibrium in the continuation in which the externally enforced component of the contract has not been altered.[4]

The effect of the renegotiation opportunity at date 7 is to constrain transfers to be "balanced," that is, satisfying

$$t \in \mathbb{R}^2_0 \equiv \{t' \in \mathbb{R}^2 \mid t'_1 + t'_2 = 0\}.$$

Thus, we simply assume that transfers are balanced and then otherwise ignore date 7. Also, as we explain later, the opportunity for ex post renegotiation implies that there is never any renegotiation surplus at date 3, so we can ignore interaction at date 3.

Much of our analysis does not depend on the details of date 2 interaction, but some of our key results concern the relation between the investment and the trade technologies, and for these we need to formally distinguish between different investment technologies. We assume that a single player makes an investment choice at date 2. This gives us two cases to consider:

[3] In Sections 4 and 5, we provide some analysis for the setting in which renegotiation is possible at date 3 but not at date 5.

[4] The generalized Nash bargaining solution has this representation. The rationality conditions identify a *contractual equilibrium*; see Watson (2006) for notes on the relation between "cooperative" and "noncooperative" approaches to modeling negotiation. The players obtain the joint value $\gamma(\theta)$ because, at the time of renegotiation, they know the state θ and can select a contract that forces the action $a^*(\theta)$, as described in the next subsection.

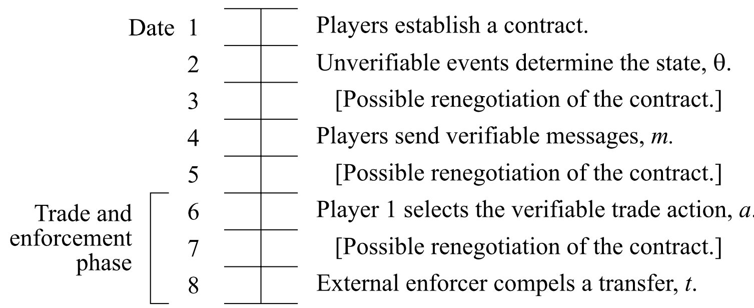

FIGURE 1. Timeline of the contractual relationship.

setting. When the state is realized, it becomes commonly known by the players; however, it cannot be verified to the external enforcer. Let Θ denote the set of possible states.

The *trade action* a. This is an individual action chosen by player 1 that determines whether and how the relationship is consummated. The trade action is commonly observed by the players and is verifiable to the external enforcer. Let A be the set of feasible trade actions.

The *monetary transfers* $t = (t_1, t_2)$. Here t_i denotes the amount given to player i for $i = 1, 2$, where a negative value represents an amount taken from this player. Transfers are compelled by the external enforcer, who is not a strategic player, but, rather, who behaves as directed by the contract of players 1 and 2.[2] Assume $t_1 + t_2 \leq 0$.

We assume that the players' payoffs are additive in money and are thus defined by a function $u\colon A \times \Theta \to \mathbb{R}^2$. In state θ, with trade action a and transfer t, the payoff vector is $u(a, \theta) + t$. Define $U(a, \theta) \equiv u_1(a, \theta) + u_2(a, \theta)$, which is the joint value of the contractual relationship in state θ if trade action a is selected. We assume that, in each state θ, the joint value has a unique maximizer $a^*(\theta)$. We let $\gamma(\theta)$ denote the maximal joint payoff in state θ, so we have

$$\gamma(\theta) \equiv U(a^*(\theta), \theta) = \max_{a \in A} U(a, \theta). \tag{1}$$

In addition to the payoff-relevant components of their relationship, we assume that the players can communicate with the external enforcer using public, verifiable messages. Let $m = (m_1, m_2)$ denote the profile of messages that the players send and let M_1 and M_2 be the sets of feasible messages. The sets M_1 and M_2 are endogenous in the sense that they are specified by the players in their contract.

Figure 1 shows the timeline of the contractual relationship. At even-numbered dates through date 6, the players make joint observations and they make individual decisions—jointly observing the state at date 2, sending verifiable messages at date 4,

[2]That the external enforcer's role is limited to compelling transfers is consistent with what courts do in practice.

- *Unified case.* Player 1 has both the date 2 investment action and the date 6 trade action.
- *Divided case.* Player 2 has the date 2 investment action, whereas player 1 has the date 6 trade action.

We assume that the investment influences the state. In the *deterministic* subcase, one of the players directly selects θ at date 2. More generally, the state may also depend on the outcome of a random variable.

Public-action modeling and forcing contracts

Because the trade action a is assumed to be taken by player 1, we have specified here an individual-action model. A public-action model, in contrast, abstracts by treating the trade action a as something that the external enforcer directly selects. Watson (2007) shows that specifying a public-action model is equivalent to examining the individual-action model but limiting attention to a particular class of contracts called *forcing contracts*, which, for any given message profile, prescribe that player 1 selects a particular trade action.

More precisely, a forcing contract specifies a large transfer from player 1 to player 2 in the event that player 1 does not take his contractually prescribed action. This transfer is sufficiently large to give player 1 the incentive to select the prescribed action in every state. Thus, the induced trade action is constant in the state, conditional on the messages sent earlier.

For example, holding the message profile fixed, the transfer function $\hat{y}$ defined as follows forces player 1 to select action $\hat{a}$ and imposes the transfer $\hat{t}$ (as though the external enforcer chose these in a public-action model):

> Let L be such that $L > \sup_{a,\theta} u_1(a,\theta) - \inf_{a,\theta} u_1(a,\theta)$. Then define $\hat{y}(\hat{a}) \equiv \hat{t}$ and, for every $a \neq \hat{a}$, set $\hat{y}(a) \equiv \hat{t} + (-L, L)$.

We use the term *forcing* for any transfer function that, given the message profile, induces player 1 to select the same trade action over all of the states.[5] We use the term *non-forcing* for transfer functions that induce player 1 to select different actions in at least two different states.

Continuation value functions

A *(state-contingent) value function* is a function from Θ to $\mathbb{R}^2$ that gives the players' expected payoff vector from the start of a given date, as a function of the state. Such a value function represents the continuation values for a given outstanding contract and equilibrium behavior. We adopt the convention of not including any sunk investment

[5] One could add a public randomization device to the model for the purpose of achieving randomization over trade actions using forcing contracts. Allowing such randomization does not expand the set of implementable value functions here.

costs from date 2 in the function u or in the representation of continuation values from later dates.

The continuation values from the start of date 3 are important to calculate, because they determine the players' incentives to invest at date 2. Thus, our chief objective is to characterize the set of *implementable value functions* from the start of date 3. A value function v is implementable if there is a contract that, if formed at date 1, would lead to continuation value $v(\theta)$ in state θ from the start of date 3 for every $\theta \in \Theta$.

Related literature

Much of the recent contract-theory literature focuses on public-action mechanism-design models. For instance, Che and Hausch (1999), Hart and Moore (1999), Maskin and Moore (1999), Segal (1999), and Segal and Whinston (2002) have basically the same setup as we do except that their models treat trade actions as public (collapsing together the trade action and the enforcement phase), so they focus on forcing contracts.[6] In some related papers, the verbal description of the contracting environment identifies individuals who take the trade actions, but the actions are effectively modeled as public due to an implicit restriction to forcing contracts. In some cases, such as with the contribution of Edlin and Reichelstein (1996), simple forcing contracts (or breach remedies) are sufficient to achieve an efficient outcome and so the restriction does not have efficiency consequences.[7]

Examples of individual-action models in the literature, among others, are the articles of Hart and Moore (1988), MacLeod and Malcomson (1993), and Nöldeke and Schmidt (1995). Also relevant is the work of Myerson (1982, 1991), whose mechanism-design analysis nicely distinguishes between inalienable individual and public actions (he uses the term "collective choice problem" to describe public-action models).

Most closely related to our work is that of Evans (2006, 2008), who emphasizes how efficient outcomes can be achieved by conditioning external enforcement on costly individual actions. Evans (2006) examines general mechanism-design problems; Evans (2008), which we discuss more in the Conclusion, examines contracting problems with specific investment and durable trading opportunities. Related as well is the work of Lyon and Rasmusen (2004), which shares the theme of Watson (2007), and the recent work of Boeckem and Schiller (2008) and Ellman (2006).[8]

[6] Aghion et al. (1994) is another example. The more recent entries by Roider (2004) and Guriev (2003) have the same basic public-action structure. Demski and Sappington (1991), Nöldeke and Schmidt (1998), and Edlin and Hermalin (2000) examine models with sequential investments in a tradeable asset; in these models, as in Maskin and Tirole (1999), transferring the asset is essentially a public action.

[7] Stremitzer (forthcoming) elaborates on Edlin and Reichelstein (1996) by examining the informational requirements of standard breach remedies (specifically, partially verifiable investments).

[8] Also related are some studies of delegation in principal–agent settings with asymmetric information, where implementable outcomes depend on whether it is the principal or the agent who has the productive action. As Beaudry and Poitevin (1995) show, ex post renegotiation imposes less of a constraint in the case of "indirect revelation" (where the agent has the productive action). Thus, if it is possible to transfer "ownership" of the productive action to the agent, the threat of ex post renegotiation provides one reason for doing this.

In classifying the related literature, another major distinction to make is between models with cross investment and models with "own investment." In the latter case, investment enhances the investing party's benefit of trade. We discuss this distinction in more detail in the next two sections. Since the holdup problem is more problematic in the case of cross investment, and there the distinction between forcing and nonforcing contracts (public- versus individual-action modeling) is critical, we concentrate on settings with significant cross investment.

2. Example

In this section, we provide a simple example of specific investment and holdup to illustrate our results. We continue to call player 1 Al and player 2 Zoe. One of the players is the *investor* at date 2 (Al in the unified case, Zoe in the divided case). The investor selects $\theta \in [0,9]$ at immediate cost $c(\theta) = 3\theta$. That is, the investor takes an action that determines the state. At date 6, Al selects a trade action $a \in [0,9]$, which we interpret as a quantity of an intermediate good that he delivers to Zoe. Payoffs are given by

$$u_1(a,\theta) = 4\sqrt{a\theta} - 4a \quad \text{and} \quad u_2(a,\theta) = 4\sqrt{a\theta}.$$

As we assume throughout, the sunk investment cost is not included in these functions and in the value functions computed below. Assume that the players have equal bargaining weights.

The joint value of the relationship in state θ is $U(a,\theta) = 8\sqrt{a\theta} - 4a$, which is maximized at $a^*(\theta) = \theta$. Therefore, the maximal joint value in state θ is $\gamma(\theta) = U(a^*(\theta),\theta) = 4\theta$. Regardless of who makes the investment, we see that the efficient level of investment maximizes $4\theta - 3\theta$. Thus the optimal investment level is $\theta^* = 9$.

Note that, for any fixed trade action, Al's and Zoe's payoffs increase equally in θ. Thus, regardless of who is the investor, this example exhibits elements of both *cross investment* and *own investment*. Own investment refers to the investment boosting the investor's gain from trade, whereas cross investment refers to increasing the gain of the other party.[9] The cross-investment element is particularly problematic, as the literature shows, because there are contingencies (typically out of equilibrium) in which the non-investing party can extract surplus from the investor by threatening to hold up trade. This can distort the investor's incentives and lead to inefficient investment.[10]

In fact, Che and Hausch (1999) conclude that with significant cross investment, the "null contract"—forcing no trade, regardless of the messages—is best. These authors formulate a public-action model, which limits attention to forcing contracts. Indeed, it is straightforward to show that the null contract is the best forcing contract for our

[9] Che and Hausch (1999) use the term "cooperative investment" for cross investment.

[10] The literature demonstrates that forcing contracts can usually prevent the holdup problem in the own-investment case, where the investing party obtains a large share of the benefit created by the investment. See, for example, Chung (1991), Rogerson (1992), Aghion et al. (1994), Nöldeke and Schmidt (1995), and Edlin and Reichelstein (1996). An exception is the "complexity/ambivalence" setting studied by Segal (1999), Hart and Moore (1999), and Reiche (2006).

Consider the following dual-option contract that is parameterized by a number $\alpha \in [0, 9]$. At date 4, Zoe must send a message $\hat{a} \geq 0$, which we interpret as a requested quantity for Al to deliver. Then Al is forced to choose between $a = \hat{a}$ and $a = \alpha$ at date 6, by having him pay Zoe a large amount if he selects any other trade action. That is, if the contract remains in place by date 8 and Al selects some $a \notin \{\alpha, \hat{a}\}$, then the enforcer compels a large transfer (say, 100) from Al to Zoe. If Al selects $a = \alpha$, then there is no transfer, whereas if Al selects $a = \hat{a}$, then the enforcer compels a transfer of $u_1(\alpha, \hat{a}) - u_1(\hat{a}, \hat{a}) = 4\sqrt{\alpha\hat{a}} - 4\alpha$ from Zoe to Al.

Let us construct a value function that this dual-option contract implements. Note that, given α and absent renegotiation at date 5, Al weakly prefers to choose $a = \alpha$ at date 6 if and only if

$$u_1(\alpha, \theta) \geq u_1(\hat{a}, \theta) + [u_1(\alpha, \hat{a}) - u_1(\hat{a}, \hat{a})].$$

Plugging in the values, this simplifies to

$$\sqrt{\theta}(\sqrt{\alpha} - \sqrt{\hat{a}}) \geq \sqrt{\hat{a}}(\sqrt{\alpha} - \sqrt{\hat{a}}).$$

Note that if Zoe requests $\hat{a} = \theta$ (the ex post optimal trade level for the realized state θ), then Al is indifferent between trade actions θ and α at date 6. Let us assume that Al selects $a = \theta$ in this contingency. Because $a = \theta$ is the efficient trade action in state θ, the players would not renegotiate at date 5, so the payoffs from date 4 are $u_1(\alpha, \theta)$ for Al and $\gamma(\theta) - u_1(\alpha, \theta)$ for Zoe.

Observe that Zoe can do no better by deviating from $\hat{a} = \theta$ at date 4. This is because Al can ensure himself a payoff of at least $u_1(\alpha, \theta)$ from date 6 by choosing $a = \alpha$. Renegotiation earlier can only add to Al's pocket, so his continuation payoff from date 4 is bounded below by $u_1(\alpha, \theta)$. Since the opportunity for renegotiation implies that Zoe's continuation value is $\gamma(\theta)$ minus Al's continuation value and since Zoe can hold Al down to $u_1(\alpha, \theta)$ by choosing $\hat{a} = \theta$, it is optimal for Zoe to send this message. Thus, in state θ, Zoe requests $\hat{a} = \theta$, there is no renegotiation, and Al delivers θ units. The contract implements the value function given by

$$v(\theta) = (u_1(\alpha, \theta), \gamma(\theta) - u_1(\alpha, \theta)) = (4\sqrt{\alpha\theta} - 4\alpha, 4\theta - 4\sqrt{\alpha\theta} + 4\alpha)$$

for all $\theta \in [0, 9]$. Incentives are the same if we include a constant transfer τ from Zoe to Al, in which case the implemented value function is given by

$$v(\theta) = (4\sqrt{\alpha\theta} - 4\alpha + \tau, 4\theta - 4\sqrt{\alpha\theta} + 4\alpha - \tau).$$

We next investigate the implications of this dual-option contract for the divided and unified cases.

Dual-option contract in the divided case

The implication of utilizing nonforcing contracts is dramatic in the divided case, where Zoe makes the investment. In fact, it is easy to see that the dual-option contract can be

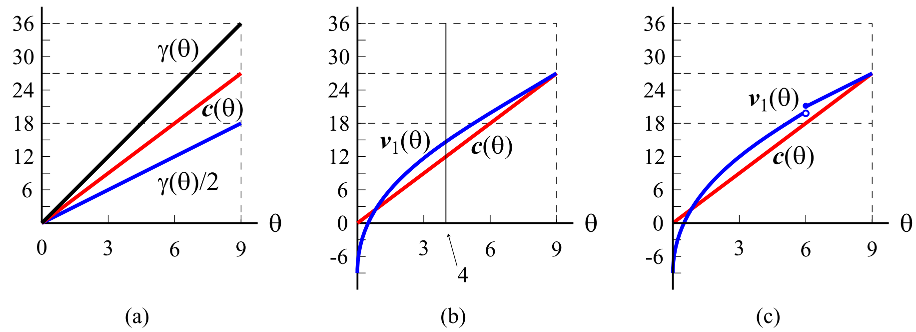

FIGURE 2. Value function and investment cost.

example, regardless of which player is the investor.[11] Unfortunately, the null contract leads to an inefficient level of investment. To see this, note that the players always renegotiate to take the ex post efficient trade action in each state. In our example this implies that in state θ, the renegotiation surplus equals the joint value 4θ. Since the investor receives half of the surplus (recall that the bargaining weights are $1/2$), the investor's value from date 3 is 2θ. This value and the investment cost are illustrated in part (a) of Figure 2. At date 2, the investor therefore chooses $\theta \in [0, 9]$ to maximize $2\theta - 3\theta$, and so the inefficiently low level $\theta = 0$ is chosen.

Implementation with a dual-option contract

We next demonstrate that by using nonforcing contracts, a more efficient outcome can be achieved (Watson's 2007 point) and that the unified and divided cases behave quite differently. Our analysis features a particular nonforcing contract that we call a *dual option*, which turns out to be an optimal contractual form in a wide range of settings. With the dual-option contract, Zoe sends a message at date 4 and then Al is forced to choose between two different trade actions at date 6, one of which depends on Zoe's message. Thus, Zoe's option is message-based, whereas Al's option later is his choice of the trade action.[12]

[11] The tools developed in Sections 4 and 5 can be used to show that if v is implemented by a forcing contract, then, for any θ, θ' with $\theta > \theta'$, $v_1(\theta) - v_1(\theta')$ and $v_2(\theta) - v_2(\theta')$ are bounded above by $2(\theta - \theta')$. The null contract achieves this bound. See Appendix B for more details.

[12] The literature emphasizes the importance of option contracts for aligning incentives. By laying out the details of the trade technology, we are able to differentiate between message-based and trade-action-based option components. Our dual-option contract is a novel addition to the theoretical literature because it has both of these components. By comparison, papers in the related literature examine either (i) one-sided or two-sided message-based contracts, or (ii) action-based options without messages. Che and Hausch (1999) were first to demonstrate the advantages of a sequential, two-sided message-based contract in a setting without renegotiation (where trade-action-based options offer no special advantage); Segal and Whinston (2002) examine general two-sided message contracts with renegotiation; Nöldeke and Schmidt (1995) look at action-based contracts without messages; Hart and Moore (1988) and MacLeod and Malcomson (1993) also have action-based contracts but with partial verifiability (which we further discuss in the Conclusion). Demski and Sappington (1991) briefly discuss a contract with both message- and trade-based option components, but they do not formally study this form.

used to make Zoe the residual claimant with respect to the post-investment joint value. In particular, set $\alpha = 0$ so that Al always has the option to deliver nothing and get no transfer. Also set $\tau = 0$. This dual-option contract implements

$$v(\theta) = (0, \gamma(\theta)) = (0, 4\theta)$$

for all θ. Zoe obtains exactly the joint value of her investment, so at date 2 she maximizes $\gamma(\theta) - c(\theta)$. Her optimal choice is $\theta^* = 9$ and thus efficient investment and trade are realized.[13]

Dual-option contract in the unified case

Next consider the unified case of the example, in which Al makes the investment. We see that this case is more problematic, but that positive investment can still be induced. Consider the dual-option contract with $\alpha = 9$ and $\tau = 27$. This contract implements

$$v_1(\theta) = u_1(9, \theta) + 27 = 12\sqrt{\theta} - 9$$

for all θ. The value function is shown in part (b) of Figure 2. At date 2, Al chooses θ to maximize $v_1(\theta) - c(\theta) = 12\sqrt{\theta} - 3\theta - 9$, and so his optimal investment choice is $\theta = 4$. Thus, the dual-option contract performs better than the null contract, but does not induce efficient investment.

Consider next a version of the dual-option contract with an additional parameter $\beta \in [0, 9]$, where (i) if Zoe's message satisfies $\hat{a} \geq \beta$, then Al is forced to deliver α units with transfer τ, and (ii) if $\hat{a} < \beta$, then the specifications are as described above. One can verify that Zoe optimally selects $\hat{a} = \theta$ as before, but now the parties renegotiate whenever $\theta \geq \beta$. Setting $\alpha = 9$ and $\tau = 27$ once again, this dual-option contract implements the value function given by

$$v_1^{\beta}(\theta) = \begin{cases} 12\sqrt{\theta} - 9 & \text{if } \theta < \beta \\ 2\theta + 9 & \text{if } \theta \geq \beta, \end{cases}$$

with $v_2^{\beta}(\theta) = \gamma(\theta) - v_1^{\beta}(\theta)$. The value function is pictured in part (c) of Figure 2.

The results we provide in the next section establish that this dual-option contract is, in fact, optimal for some β. The advantage of the cutoff β is that v_1 jumps up at this point (owing to the positive renegotiation surplus), giving player 1 an extra incentive to invest at level β. The highest investment level supported by this type of contract is $\theta = 6$, which is achieved by setting $\beta = 6$. An implication is that the efficient investment level cannot be achieved.

[13] One perspective on the related literature comes from considering how contractual elements act to effectively shift relative bargaining power away from the intrinsic level inherent in the exogenous bargaining protocol. The dual option here produces the same outcome that arises under a contract that forces no trade, if one were able to give Zoe all of the bargaining power in renegotiation. Thus, the action-based part of the dual option sets a reservation payoff level for Al that allows Zoe to extract the full surplus by sending the appropriate message. A similar type of intuition is at play in Aghion et al.'s (1994) use of a financial bond to effectively shift bargaining weights, and it is also present in MacLeod and Malcomson's (1993) use of an outside option in the context of ongoing renegotiation with a durable trading opportunity.

In summary, the example shows that by using nonforcing contracts, the parties can achieve a more efficient outcome than is possible with forcing contracts. Furthermore, the efficiency gain depends on the relation between the technology of trade and the technology of investment. In the divided case, the dual-option contract induces efficient investment and trade actions. In the unified case, the efficient outcome cannot be attained, but a nonforcing contract still is preferred.

3. Investment incentives and residual claimancy

In this section we provide general versions of the results shown in the example. We divide the analysis into two subsections, one dealing with the objective of giving investment incentives to player 2 (which is needed in the divided case) and one with the objective of giving such incentives to player 1 (for the unified case).

Before proceeding, it is useful to define some additional notation. Regardless of which player has the investment choice at date 2, let the investment be denoted $x \geq 0$. We normalize the investment variable so that the immediate cost of investment is exactly x for the investor. The state θ is then drawn from a distribution $G(x)$ that depends on the investment choice.[14] Recalling that $\gamma(\theta) = U(a^*(\theta), \theta)$ is the maximum joint value in state θ, we see that the efficient level of investment x^* solves

$$\max_{x \geq 0} \int \gamma(\theta)\, dG(x) - x.$$

In the *deterministic* case in which there is no random element, we suppose that $\theta \equiv x$ and so the investor selects θ directly; then we write θ^* as the efficient investment choice, which maximizes $\gamma(\theta) - \theta$.

Letting i denote the investing party, we want to implement a value function v so that $v_i(\theta)$ is increasing in θ to some particular extent. In this way, player i is rewarded for investing. Ideally, it would be possible to implement a value function that satisfies $v_i(\theta) = \gamma(\theta) - k$ for all $\theta \in \Theta$ and some fixed k, because this makes player i the residual claimant with respect to his/her investment decision. Player i's payoff from date 2 is then

$$\int v_i(\theta)\, dG(x) - x = \int \gamma(\theta)\, dG(x) - x - k,$$

and so player i optimally selects x^*, leading to efficient investment and trade. The players select k to divide the joint value at date 1.

It may also be possible to achieve an efficient outcome without having $v_i = \gamma - k$, but this is not always the case. More generally, in some settings we can characterize the optimal contract and the best (though inefficient) investment that can be induced. Our proofs utilize a dual-option contract and thus further demonstrate the utility of this simple contractual form.

[14]It is natural to assume that G is increasing in x in the sense of first-order stochastic dominance and that $U(a^*(\theta), \theta)$ is increasing in θ, so that higher investments increase the expected gains from trade, but these assumptions are not needed for Proposition 1 below.

Investment incentives for player 2

We start by showing that, assuming that investment conveys no instantaneous (direct) benefits, a dual-option contract can be used to make player 1's value from date 3 constant in the state. Player 2 then becomes the residual claimant. Thus, in the divided case in which player 2 is the investor, player 2 can be given the incentive to invest efficiently regardless of the distribution of the investment gains.

ASSUMPTION 1. *There exists a trade action $a^0 \in A$ such that $u_1(a^0, \theta) = u_2(a^0, \theta) = 0$ for every θ.*

This assumption is solely on the technology of trade. Think of a^0 as the "no trade" choice. The no-trade payoffs could be normalized to any level; we set them to zero here for simplicity. Note that the example satisfies Assumption 1 with $a^0 = 0$.

THEOREM 1. *Consider any contractual relationship that satisfies Assumption 1. Let k be any real number and define the value function v by $v_1(\theta) = k$ and $v_2(\theta) = \gamma(\theta) - k$ for all $\theta \in \Theta$. Then v is implementable.*

Appendix A contains the proof of this theorem, which is constructive and runs along the lines of the demonstration for the example. In particular, we show how to implement these value functions using a dual-option contract in which player 2 is required to declare a state $\hat{\theta}$ at date 4 and then player 1 is forced to tender either trade action $a^*(\hat{\theta})$ or trade action a^0.

Considering the implication of making the investor the residual claimant, Theorem 1 leads immediately to the following economic result.

PROPOSITION 1. *Under Assumption 1 and in the divided case in which player 2 is the investor and player 1 has the trade action, optimal contracting induces efficient investment and trade (the first-best outcome).*

Note that this result makes no restrictions on which party stands to gain from the investment. That is, the result holds for settings of cross investment, own investment, and any combination of the two. The key is simply that the investment action and the trade action are taken by different parties.

Investment incentives for player 1

Making player 1 the residual claimant with respect to the investment decision is considerably more difficult than is making player 2 the residual claimant. In fact, in the case of unified investment and trade actions, typically the efficient outcome cannot be induced. Identifying an optimal contract is also a challenge, but is possible with some additional assumptions and structure on the model. We present results along these lines below. The proofs require a full analysis of the conditions for implementation, which in turn requires the tools and results developed in Sections 4 and 5. Details of the analysis are provided in Appendix B, where we also present some general results and notes.

We begin with some intuition regarding the conditions for implementation. Suppose that γ is increasing. Our objective is to implement a value function v_1 that rises with γ, so that player 1's return on investment closely matches the joint return. The technical conditions for implementation imply upper bounds on the difference $v_1(\theta) - v_1(\theta')$ for $\theta > \theta'$. Observe that there are multiple such conditions involving each state. For example, for three states θ, θ', and θ'' with $\theta > \theta' > \theta''$, there are three conditions: $v_1(\theta) - v_1(\theta') \le \rho$, $v_1(\theta') - v_1(\theta'') \le \rho'$, and $v_1(\theta) - v_1(\theta'') \le \rho''$ for some numbers ρ, ρ', and ρ''. We can call the first and second conditions *local*, or *inside*, conditions, whereas the last one is an *outside* condition. Note that by summing the inside conditions, we obtain a second bound on the difference $v_1(\theta) - v_1(\theta'')$; this bound is $\rho + \rho'$.

It turns out that, for a wide class of trade technologies, the outside condition is tighter than is the sum of the inside conditions; that is, $\rho'' < \rho + \rho'$. This means that implementability cannot be characterized by the local conditions alone, and some of these must hold with slack to ensure that the outside conditions are satisfied. As a result, it is not possible to implement a value function v_1 that rises smoothly and steeply. This is not such a big problem in the divided case, where we want v_1 to be constant, but recall that in the unified case we want v_1 to rise with γ. We find that the best way to give player 1 the incentive to invest is to implement a value function with some discrete jumps. We establish conditions under which a dual-option contract optimally performs in this way, as shown in the example.

We make several assumptions to structure the analysis. The first gives a set of mild technical restrictions that hold in most applications. We maintain this assumption throughout the rest of the paper.

Assumption 2. *(a) The sets A and Θ are compact subsets of $\mathbb{R}$ and contain at least two elements, and $u_1(\cdot, \theta)$ and $u_2(\cdot, \theta)$ are continuous functions of a for every $\theta \in \Theta$. Define $\underline{a} \equiv \min A$, and $\overline{a} \equiv \max A$, $\underline{\theta} \equiv \min \Theta$, and $\overline{\theta} \equiv \max \Theta$. (b) $U(\cdot, \theta)$ is strictly quasiconcave for every $\theta \in \Theta$. (c) Player 1's bargaining weight is positive: $\pi_1 > 0$.*

The next assumption is the main economic restriction that we impose hereinafter: that player 1's payoff is supermodular in the state and trade action.

Assumption 3. *The function u_1 is supermodular, meaning that $u_1(a, \theta) - u_1(a', \theta) \ge u_1(a, \theta') - u_1(a', \theta')$ whenever $a \ge a'$ and $\theta \ge \theta'$.*

With this assumption, player 1's marginal value of increasing his trade action rises weakly with the state. In other words, higher trade actions are weakly more attractive to him as the state increases. An implication is that, for any transfers specified as a function of the trade action, player 1's preferences satisfy the single-crossing property and he weakly prefers higher actions in higher states. This monotone structure helps us to characterize incentives at date 6.

Many interesting applications studied in the literature satisfy these assumptions. For instance, consider a buyer/seller relationship in which a is the number of units of an intermediate good to be transferred from the seller to the buyer. The buyer's benefit of obtaining a units in state θ is $B(a,\theta)$. The seller's cost of production and delivery is $d(a,\theta)$, and we let $C(a,\theta) = -d(a,\theta)$. Suppose, as one typically does, that B is increasing and concave in a and that d is increasing and convex in a. If a is the buyer's action (he selects how many units to install, for example), then the buyer is player 1 and so we have $u_1 \equiv B$ and $u_2 \equiv C$. If the seller chooses a (she decides how many units to deliver, say), then the seller is player 1 and so we have $u_1 \equiv C$ and $u_2 \equiv B$. In either case, Assumption 2 is satisfied. Assumption 3 adds the weak supermodularity requirement on the payoff of the player who selects a. Our example from the previous section satisfies Assumptions 2 and 3.

Note that, in a given application, if u_1 is submodular, then one can redefine the trade action to be $-a$ and then Assumption 3 is satisfied. Also note that Assumption 3 is trivially satisfied in the case of *pure cross investment* in which u_1 does not depend on θ.

Our final assumption, which is needed only for the results in this subsection, pertains to the relative supermodularity and investment returns for u_1 and u_2.

ASSUMPTION 4. *(a) The expression $\pi_2 u_1 - \pi_1 u_2$ is supermodular ($\pi_2 u_1$ is relatively more supermodular than $\pi_1 u_2$). (b) For all θ, θ' with $\theta > \theta'$, $\pi_1[u_2(\overline{a},\theta) - u_2(\overline{a},\theta')] \geq \pi_2[u_1(\overline{a},\theta) - u_1(\overline{a},\theta')]$.*

Assumption 4 is clearly restrictive, limiting the class of trade technologies that we evaluate here, but it facilitates the identification of an optimal contract. Part (a) of the assumption contributes to a tight characterization of optimal punishments in the general mechanism-design exercise. A sufficient condition for Assumption 4(a) is that u_2 is submodular. Appendix B gives an alternative assumption—on the joint value at extreme trade actions—that can be used in place of Assumption 4(a). Assumption 4(b) requires that the cross-investment component is weakly larger than the own-investment component at the highest trade action. It is easy to check that our example satisfies Assumption 4.

These assumptions give us the following general result about implementation using the dual-option contract introduced in the example.

THEOREM 2. *Consider any contractual relationship that satisfies Assumptions 2–4. For any number β, define value function v^β by*

$$v_1^\beta(\theta) \equiv \begin{cases} u_1(\overline{a},\theta) & \text{for } \theta < \beta \\ u_1(\overline{a},\theta) + \pi_1 R(\overline{a},\theta) & \text{for } \theta \geq \beta \end{cases}$$

and $v_2^\beta \equiv \gamma - v_1^\beta$ for all $\theta \in \Theta$. Then v^β is implemented by a dual-option contract in which (i) at date 4, player 2 sends a report $\hat{\theta}$ of the state; and (ii) if the report is at least β, then player 1 is forced to select $a = \overline{a}$ at date 6, and otherwise player 1 is forced to choose between $a^*(\hat{\theta})$ and $\overline{a}$.

The proof of Theorem 2 is provided in Appendix B, along with the proofs of Propositions 2 and 3 below. Note that we have not used Assumption 1 here.

We next show that value function v^{β} achieves the best possible investment incentives in the deterministic case where player 1's investment choice is to directly select θ. Thus, in this setting of unified investment and trade actions, the dual-option contract is optimal. Recall that we normalize so that the cost of investment is θ, and thus player 1 selects θ to maximize $v_1(\theta) - \theta$. The efficient choice θ^* maximizes $\gamma(\theta) - \theta$.

Let us say that *investment* θ' *is supported* by a contract if there is an implementable value function v' such that θ' solves $\max_{\theta\in\Theta} v_1'(\theta) - \theta$. Call θ^{B} the *best achievable* investment level if it maximizes $\gamma(\theta') - \theta'$ among all supportable θ'.

Proposition 2. *Under Assumptions 2–4, and in the deterministic and unified case in which player 1 has both the investment and trade actions, the best achievable investment level* θ^{B} *is supported by value function* $v^{\theta^{\mathrm{B}}}$ *as defined in Theorem 2 (that is, setting* $\beta \equiv \theta^{\mathrm{B}}$*).*

Our next result gives conditions under which the efficient investment level can be supported; that is, when the best achievable investment level θ^{B} coincides with the efficient investment level θ^*.

Proposition 3. *Suppose Assumptions 2–4 hold. The efficient level of investment* θ^* *is supported in the deterministic and unified case if and only if*

$$u_1(\overline{a}, \theta^*) + \pi_1 R(\overline{a}, \theta^*) - \theta^* \geq u_1(\overline{a}, \theta) - \theta \tag{2}$$

for all $\theta < \theta^*$.

The condition from Proposition 3 ensures that we can induce a large enough discontinuity in the value function at the efficient state so that player 1 maximizes his gain net of investment cost at θ^*. If this condition fails, efficiency cannot be attained in the unified case.

To summarize the results of this section, we have conditions under which player 2 can be made the residual claimant with respect to the investment action, which solves the contracting problem (yielding efficient investment and trade) in the divided case. We learn that it is generally not possible to make player 1 the residual claimant, given that player 1 has the trade action. As a result, the efficient outcome is typically not attainable in the case of unified investment and trade actions. However, for the class of trade technologies that satisfy Assumptions 2–4 and for the case of a deterministic state, we are able to characterize an optimal contract and provide conditions under which the efficient outcome can be achieved.

The results in this section are most pronounced when applied to settings of significant cross investment, where the optimal forcing contract is null and leads to an inefficient outcome. In the divided case, the nonforcing dual-option contract induces efficient investment and trade. In the unified case, a dual-option contract can outperform the null contract and sometimes induces the efficient outcome. We continue this theme in Section 5 by simply asking whether, in general, nonforcing contracts implement a wider range of value functions than do forcing contracts.

4. Implementable value functions

This section summarizes how to calculate implementable value functions in general. Much of the analysis here repeats material in Watson (2007), so we keep this text brief and ask the reader to see Watson (2007) for more details. The culmination of the basic analysis here are some simple characterization results from Watson (2007), which we build on in the subsequent section.

In previous sections we assumed that the players can freely renegotiate at dates 3 and 5, but now we also consider the case in which renegotiation is possible at date 3 only (the interim phase). We let V^{EPF} be the set of implementable value functions from date 3 for the case of ex post renegotiation and with the restriction to forcing contracts. We let V^{EP} be the corresponding set for the case of ex post renegotiation and no contractual restrictions. Further, we let V^{I} be the set of implementable value functions for the case in which renegotiation can occur only at date 3.[15] We can characterize the implementable value functions by backward induction, starting with date 6 where player 1 selects the trade action.

State-contingent values from date 6

To calculate the value functions that are supported from date 6, we can ignore the payoff-irrelevant messages sent earlier (or equivalently, fix a message profile from date 4) and simply write the externally enforced transfer function as $\hat{y}: A \to \mathbb{R}^2$. That is, $\hat{y}$ gives the monetary transfer as a function of player 1's trade action.

Given the state θ, $\hat{y}$ defines a *trading game* in which player 1 selects an action $a \in A$ and the payoff vector is then $u(a,\theta) + \hat{y}(a)$. Focusing on pure strategies, we let $\hat{a}(\theta)$ denote the action chosen by player 1 in state θ. This specification is rational for player 1 if, for every $\theta \in \Theta$, $\hat{a}$ maximizes $u_1(a,\theta) + \hat{y}_1(a)$ by choice of a. The state-contingent payoff vector from date 6 is then given by the *outcome function* $w: \Theta \to \mathbb{R}^2$ defined by

$$w(\theta) \equiv u(\hat{a}(\theta), \theta) + \hat{y}(\hat{a}(\theta)). \tag{3}$$

Let W denote the set of supportable outcome functions. That is, $w \in W$ if and only if there are functions $\hat{y}$ and $\hat{a}$ such that $\hat{a}$ is rational for player 1 and, for every $\theta \in \Theta$, (3) holds. Furthermore, let W^{F} be the subset of outcomes that can be supported using forcing contracts. It is easy to see that $w \in W^{\mathrm{F}}$ if and only if there is a trade action $\hat{a}$ and a transfer vector $\hat{t}$ such that $w(\theta) = u(\hat{a},\theta) + \hat{t}$ for all $\theta \in \Theta$. We can compare individual-action and public-action models by determining whether the restriction to forcing contracts implies a significant constraint on the set of implementable value functions.

State-contingent values from date 5

We next step back to date 5. If there is no opportunity for ex post renegotiation, then nothing happens at date 5 and so W and W^{F} are the supported state-contingent value

[15] In the case of only interim renegotiation, a restriction to forcing contracts does not affect the implementable set.

sets from the start of date 5 as well. On the other hand, if ex post renegotiation is allowed, then at date 5 the players have an opportunity to discard their originally specified contract y and replace it with another, y'.

By picking a new contract y', the players are effectively choosing a new outcome function w' in place of the function w that would have resulted from the original contract y. The players can freely select w' from the set W or the set W^{F}, depending on whether they are restricted to forcing contracts. The players divide the renegotiation surplus according to the fixed bargaining weights π_1 and π_2. Dividing the surplus in this way is feasible because W and W^{F} are closed under constant transfers.

Clearly, we have $\gamma(\theta) = \max_{w \in W^{\mathrm{F}}}[w_1(\theta) + w_2(\theta)]$ because the trade action that solves the maximization problem in (1) can be specified in a forcing contract to yield the desired outcome. The *renegotiation surplus* is

$$r(w, \theta) \equiv \gamma(\theta) - w_1(\theta) - w_2(\theta).$$

The bargaining solution implies that the players settle on a new outcome in which the payoff vector in state θ is $w(\theta) + \pi r(w, \theta)$.

We define an *ex post renegotiation outcome* to be the state-contingent payoff vector that results when, in every state, the players renegotiate from a fixed outcome in W. That is, a value function z is an ex post renegotiation outcome if and only if there is an outcome $w \in W$ such that $z(\theta) = w(\theta) + \pi r(w, \theta)$ for every $\theta \in \Theta$. Let Z denote the set of ex post renegotiation outcomes.[16] If trade actions are treated as public (and so attention is limited to forcing contracts), then the set of ex post renegotiation outcomes contains only the value functions of the form $z = w + \pi r(w, \cdot)$ with the constraint that $w \in W^{\mathrm{F}}$. Let Z^{F} denote the set of ex post renegotiation outcomes under forcing contracts. Although the terminology is a bit loose, we refer to functions in Z and Z^{F}, in addition to functions in W and W^{F}, simply as "outcomes."

State-contingent values from dates 4 and 3

Analysis of contract selection and incentives at date 4 can be viewed as a standard mechanism-design problem. The players' contract is equivalent to a mechanism that maps messages sent at date 4 to outcomes induced in the trade and enforcement phase (possibly renegotiated at date 5). The revelation principle applies, so we can restrict attention to direct-revelation mechanisms defined by (i) a message space $M \equiv \Theta^2$ and (ii) a function that maps Θ^2 to the relevant outcome set that gives the state-contingent value functions from the start of date 5. The outcome set is either W, W^{F}, Z, or Z^{F}, depending on whether ex post renegotiation and/or nonforcing contracts are allowed. We concentrate on Nash equilibria of the mechanism in which the parties report truthfully in each state.[17]

Let us write $\psi^{\theta_1\theta_2}$ for the outcome that the mechanism prescribes when player 1 reports the state to be θ_1 and player 2 reports the state to be θ_2. Note that, in any given

[16] All elements of Z are efficient in every state; also, Z and W are generally not ranked by inclusion.

[17] The revelation principle usually requires a public randomization device to create lotteries over outcomes (or that the outcome set is a mixture space), but it is not needed here.

state θ (the actual state that occurred), the mechanism implies a "message game" with strategy space Θ^2 and payoffs given by $\psi^{\theta_1\theta_2}(\theta)$ for each strategy profile (θ_1, θ_2). For truthful reporting to be a Nash equilibrium of this game, it must be that $\psi_1^{\theta\theta}(\theta) \geq \psi_1^{\tilde{\theta}\theta}(\theta)$ and $\psi_2^{\theta\theta}(\theta) \geq \psi_2^{\theta\tilde{\theta}}(\theta)$ for all $\tilde{\theta} \in \Theta$.

We proceed using standard techniques for mechanism design with transfers, following Watson (2007). The key step is observing that, for any two states θ and θ', the outcome specified for the "off-diagonal" message profile (θ', θ) must be sufficient to simultaneously (i) dissuade player 1 from declaring the state to be θ' when the state is actually θ and (ii) discourage player 2 from declaring θ in state θ'. Thus, we require

$$\psi_1^{\theta\theta}(\theta) \geq \psi_1^{\theta'\theta}(\theta) \quad \text{and} \quad \psi_2^{\theta'\theta'}(\theta') \geq \psi_2^{\theta'\theta}(\theta').$$

Because the outcome sets are closed under constant transfers, we can choose the outcome to effectively raise or lower $\psi_1^{\theta'\theta}$ and $\psi_2^{\theta'\theta}$ while keeping the sum constant. Thus, a sufficient condition for these two inequalities is that the sum of the two holds. Letting $\psi \equiv \psi^{\theta\theta}$ and $\psi' \equiv \psi^{\theta'\theta'}$, we thus have the following necessary condition for implementing outcome ψ in state θ and outcome ψ' in state θ':

$$\text{There exists an outcome } \hat{\psi} \text{ satisfying } \psi_1(\theta) + \psi_2'(\theta') \geq \hat{\psi}_1(\theta) + \hat{\psi}_2(\theta').$$

This condition, applied to all ordered pairs (θ, θ'), is necessary and sufficient for implementation. The sum $\hat{\psi}_1(\theta) + \hat{\psi}_2(\theta')$ is called the *punishment value* corresponding to the ordered pair (θ, θ'). The punishment value plays a central role in our analysis. Lower punishment values imply a greater set of implementable outcomes.

Interim renegotiation has the effect of requiring each "on-diagonal" outcome to be efficient in the relevant state; that is, for each θ, we need $\psi^{\theta\theta}$ to be efficient in this state. In the case of ex post renegotiation, allowing interim renegotiation entails no further constraint because every outcome in Z is efficient in every state. It is also the case that without ex post renegotiation, W and W^{F} yield the same set of implementable value functions from date 3. Therefore, we have three settings to compare: unrestricted contracts with ex post renegotiation, forcing contracts (public actions) with ex post renegotiation, and forcing contracts with interim (but not ex post) renegotiation.

Call a value function v *efficient* if $v_1(\theta) + v_2(\theta) = \gamma(\theta)$ for every $\theta \in \Theta$. The following results summarize the characterization of V^{EP}, V^{EPF}, and V^{I} and provide a general comparison.

RESULT 1 (Watson (2007)). Consider any value function $v : \Theta \to \mathbb{R}^2$.

- *Implementation with interim renegotiation.* Value function v is an element of V^{I} if and only if v is efficient and, for every pair of states θ and θ', there is an outcome $\hat{w} \in W^{\mathrm{F}}$ such that $v_1(\theta) + v_2(\theta') \geq \hat{w}_1(\theta) + \hat{w}_2(\theta')$.

- *Implementation with ex post renegotiation.* Value function v is an element of V^{EP} if and only if v is efficient and, for every pair of states θ and θ', there is an outcome $\hat{z} \in Z$ such that $v_1(\theta) + v_2(\theta') \geq \hat{z}_1(\theta) + \hat{z}_2(\theta')$.

- *Implementation with ex post renegotiation and forcing contracts.* Value function v is an element of V^{EPF} if and only if v is efficient and, for every pair of states θ and θ', there is an outcome $\hat{z} \in Z^{\mathrm{F}}$ such that $v_1(\theta) + v_2(\theta') \geq \hat{z}_1(\theta) + \hat{z}_2(\theta')$.

Furthermore, the sets V^{EP}, V^{EPF}, and V^{I} are closed under constant transfers.

Result 2 (Watson (2007)). The implementable sets are weakly nested in that $V^{\mathrm{EPF}} \subseteq V^{\mathrm{EP}} \subseteq V^{\mathrm{I}}$. Furthermore, $V^{\mathrm{EPF}} = V^{\mathrm{EP}}$ if and only if, for every pair of states $\theta, \theta' \in \Theta$ and every $\hat{z} \in Z$, there is an ex post renegotiation outcome $\tilde{z} \in Z^{\mathrm{F}}$ such that $\tilde{z}_1(\theta) + \tilde{z}_2(\theta') \leq \hat{z}_1(\theta) + \hat{z}_2(\theta')$. Likewise, $V^{\mathrm{EP}} = V^{\mathrm{I}}$ if and only if, for all $\theta, \theta' \in \Theta$ and every $\hat{w} \in W^{\mathrm{F}}$, there is an ex post renegotiation outcome $\hat{z} \in Z$ such that $\hat{z}_1(\theta) + \hat{z}_2(\theta') \leq \hat{w}_1(\theta) + \hat{w}_2(\theta')$.[18]

To summarize, we have thus far analyzed the players' behavior at the various dates in the contractual relationship, leading to a simple characterization of implementable value functions from date 3. The characterization is in terms of the minimum punishment values for each pair of states, which yields a way to relate the implementable sets for the cases of interim renegotiation, ex post renegotiation, and ex post renegotiation and forcing contracts. We next turn to investigate the relation more deeply.

5. A robustness result for non-forcing contracts

The example from Watson (2007) and our example in Section 2 provide illustrations of $V^{\mathrm{EPF}} \neq V^{\mathrm{EP}} \neq V^{\mathrm{I}}$. Our main objective in this section is to examine the robustness of this conclusion. We consider the wide class of contractual relationships that satisfy Assumptions 2, 3, and 5 (which follows).

Assumption 5. *There exist states $\theta^1, \theta^2 \in \Theta$ such that $\theta^1 > \theta^2$, $a^*(\theta^1) > \underline{a}$, $a^*(\theta^2) < \overline{a}$ and either $U(\underline{a}, \theta^2) < U(\overline{a}, \theta^2)$ or $U(\underline{a}, \theta^1) > U(\overline{a}, \theta^1)$.*

This is a weak assumption that removes a knife-edge case concerning the relative joint values of the extreme trade actions in the various states. For instance, if Θ has more than two elements and $U(\underline{a}, \theta) \neq U(\overline{a}, \theta)$ for some θ strictly between $\underline{\theta}$ and $\overline{\theta}$ with interior optimal actions, then Assumption 5 is satisfied.

We have the following robustness result.

Theorem 3. *Consider any contractual relationship that satisfies Assumptions 2, 3, and 5. The sets of implementable value functions in the cases of unrestricted contracts with ex post renegotiation, forcing contracts with ex post renegotiation, and interim renegotiation are all distinct. That is, $V^{\mathrm{EPF}} \neq V^{\mathrm{EP}} \neq V^{\mathrm{I}}$.*

[18] Watson's (2005) Lemma 1 provides some of the supporting analysis (which was not explained fully in the relevant proof in Watson 2007). This lemma establishes that, for any given ordered pair of states θ and θ' and any supportable outcome ψ, there exists an implementable value function v for which $v_1(\theta) + v_2(\theta') = \psi_1(\theta) + \psi_2(\theta')$. Because the minimum punishment values exist, in each case we can let ψ equal the outcome that attains the minimum.

The analysis underlying Theorem 3 amounts to characterizing and comparing the minimum punishment values that can be supported for each of the settings of interest. Recall that the punishment value for the ordered pair (θ, θ') is the value $\psi_1(\theta) + \psi_2(\theta')$, where ψ is the outcome specified in the message game when player 1 reports the state to be θ' and player 2 reports the state to be θ. Lower punishment values serve to relax incentive conditions, so to characterize the sets of implementable value functions completely, we must find the minimum punishment values. We let P^{I}, P^{EP}, and P^{EPF} denote the minimum punishment values for the settings of interim renegotiation, ex post renegotiation, and ex post renegotiation and forcing contracts, respectively:

$$P^{\mathrm{I}}(\theta, \theta') \equiv \min_{w \in W^{\mathrm{F}}} w_1(\theta) + w_2(\theta')$$

$$P^{\mathrm{EP}}(\theta, \theta') \equiv \min_{\hat{z} \in Z} \hat{z}_1(\theta) + \hat{z}_2(\theta')$$

$$P^{\mathrm{EPF}}(\theta, \theta') \equiv \min_{\hat{z} \in Z^{\mathrm{F}}} \hat{z}_1(\theta) + \hat{z}_2(\theta').$$

Our assumptions on the trade technology guarantee that these minima exist.

From Result 2, we know that Theorem 3 is equivalent to saying that there exist states $\theta, \theta' \in \Theta$ such that $P^{\mathrm{I}}(\theta, \theta') < P^{\mathrm{EP}}(\theta, \theta')$ and there exist (possibly different) states $\theta, \theta' \in \Theta$ such that $P^{\mathrm{EP}}(\theta, \theta') < P^{\mathrm{EPF}}(\theta, \theta')$. Thus, to prove Theorem 3, we examine the punishment values achieved by various contractual specifications in the different settings. We develop some elements of the proof in the remainder of this section; Appendix C contains the rest of the analysis. We focus in this section on the relation between V^{EPF} and V^{EP}. The analysis of the relation between V^{EP} and V^{I} is considerably simpler and is wholly contained in Appendix C.

We establish $P^{\mathrm{EP}} < P^{\mathrm{EPF}}$ by comparing the punishment values implied by (i) the outcome in which player 1 is forced to take the trade action that yields the lowest punishment value among forcing contracts, and (ii) a related nonforcing specification in which player 1 is given the incentive to select some action a in state θ and a different action a' in state θ'. We derive conditions under which a and a' can be arranged to strictly lower the punishment value for (θ, θ'), relative to the best forcing case. We then find states θ^1 and θ^2 such that the conditions must hold for at least one of the ordered pairs (θ^1, θ^2) and (θ^2, θ^1).

To explore the possible outcomes in the cases of ex post renegotiation, consider player 1's incentives at date 6. For any given transfer function $\hat{y}$, necessary conditions for player 1 to select trade action a in state θ and action a' in state θ' are

$$\begin{aligned} u_1(a, \theta) + \hat{y}_1(a) &\geq u_1(a', \theta) + \hat{y}_1(a') \\ u_1(a', \theta') + \hat{y}_1(a') &\geq u_1(a, \theta') + \hat{y}_1(a). \end{aligned} \tag{4}$$

Transfer function $\hat{y}$ can be specified so that player 1 is harshly punished for selecting any trade action other than a or a'. Then, in every state, either a or a' maximizes player 1's payoff from date 6. Thus, we can state the following fact.

Fact 1. Consider two states $\theta, \theta' \in \Theta$ and two trade actions $a, a' \in A$. Expression (4) is necessary and sufficient for the existence of a transfer function $\hat{y}: A \to \mathbb{R}^2_0$ (defined over all trade actions) such that player 1's optimal trade action in state θ is a and player 1's optimal trade action in state θ' is a'.

Summing the inequalities of expression (4), we see that there are values $\hat{y}(a), \hat{y}(a') \in \mathbb{R}^2_0$ that satisfy (4) if and only if

$$u_1(a,\theta) - u_1(a',\theta) \geq u_1(a,\theta') - u_1(a',\theta'). \tag{5}$$

Assumption 3 then implies the following fact.

Fact 2. If $\theta > \theta'$, then $a \geq a'$ implies inequality (5). If $\theta < \theta'$, then $a \leq a'$ implies inequality (5).

Note that Fact 2 gives sufficient conditions. In the case in which $u_1(\cdot,\cdot)$ is strictly supermodular (replacing weak inequalities in Assumption 3 with strict inequalities), player 1 can be given only the incentive to choose greater trade actions in higher states.

For any two states $\theta, \theta' \in \Theta$, define

$$E(\theta,\theta') \equiv \{(a,a') \in A \times A \mid \text{inequality (5) is satisfied}\}.$$

Also, for states $\theta, \theta' \in \Theta$ and trade actions $a, a' \in A$ with $(a,a') \in E(\theta,\theta')$, define

$$Y(a,a',\theta,\theta') \equiv \{\hat{y}: A \to \mathbb{R}^2_0 \mid \text{condition (4) is satisfied}\}.$$

Condition (4), combined with the identity $\hat{y}_1 = -\hat{y}_2$, implies the next fact.

Fact 3. For any $\theta, \theta' \in \Theta$ and $a, a' \in A$, with $(a,a') \in E(\theta,\theta')$, we have

$$\min_{\hat{y} \in Y(a,a',\theta,\theta')} \hat{y}_1(a) + \hat{y}_2(a') = u_1(a',\theta) - u_1(a,\theta).$$

Using the definition of the set W (recall expression (3)), any given $w \in W$ can be written in terms of the trade actions and transfers that support it. We have

$$w(\theta) = u(\hat{a}(\theta),\theta) + \hat{y}(\hat{a}(\theta))$$

and

$$w(\theta') = u(\hat{a}(\theta'),\theta') + \hat{y}(\hat{a}(\theta')),$$

where $\hat{a}$ gives player 1's choice of trade action as a function of the state and $\hat{y}$ is the transfer function that supports w.

For any state $\tilde{\theta}$ and trade action $\tilde{a}$, define $R(\tilde{a},\tilde{\theta})$ to be the renegotiation surplus if, without renegotiation, player 1 would select $\tilde{a}$. That is, $R(\tilde{a},\tilde{\theta}) = U(a^*(\tilde{\theta}),\tilde{\theta}) - U(\tilde{a},\tilde{\theta})$. Combining the expressions for w in the previous paragraph with Fact 1 and the definition of ex post renegotiation outcomes, we obtain Fact 4.

FACT 4. Consider any two states $\theta, \theta' \in \Theta$ and let α be any number. There is an ex post renegotiation outcome $z \in Z$ that satisfies $z_1(\theta) + z_2(\theta') = \rho$ if and only if there are trade actions $a, a' \in A$ and a transfer function $\hat{y}$ such that $(a, a') \in E(\theta, \theta')$, $\hat{y} \in Y(a, a', \theta, \theta')$, and

$$\rho = u_1(a, \theta) + \hat{y}_1(a) + \pi_1 R(a, \theta) + u_2(a', \theta') + \hat{y}_2(a') + \pi_2 R(a', \theta'). \tag{6}$$

In the last line, the first three terms are $w_1(\theta)$ plus player 1's share of the renegotiation surplus in state θ, totaling $z_1(\theta)$. The last three terms are $w_2(\theta')$ plus player 2's share of the renegotiation surplus in state θ', totaling $z_2(\theta')$.

Finding the best (minimum) punishment value for states θ and θ' means minimizing $\hat{z}_1(\theta) + \hat{z}_2(\theta')$ by choice of $\hat{z} \in Z$. For now, holding fixed the trade actions a and a' that player 1 is induced to select in states θ and θ', let us minimize the punishment value by choice of $\hat{y} \in Y(a, a', \theta, \theta')$. To this end, we can use Fact 3 to substitute for $\hat{y}_1(a) + \hat{y}_2(a')$ in expression (6). This yields the punishment value for trade actions a and a' in states θ and θ', respectively, written

$$\lambda(a, a', \theta, \theta') \equiv u_1(a', \theta) + \pi_1 R(a, \theta) + u_2(a', \theta') + \pi_2 R(a', \theta'). \tag{7}$$

Next, we consider the step of minimizing the punishment value by choice of the trade actions a and a', which gives us a useful characterization of $P^{\mathrm{EP}}(\theta, \theta')$. Assumption 2(a) guarantees that $\lambda(a, a', \theta, \theta')$ has a minimum.

FACT 5. The minimum punishment value in the setting of ex post renegotiation is characterized as

$$P^{\mathrm{EP}}(\theta, \theta') = \min_{(a,a') \in E(\theta,\theta')} \lambda(a, a', \theta, \theta').$$

We obtain a similar characterization of the minimal punishment value for the setting in which attention is restricted to forcing contracts. The characterization is exactly as in Fact 5 except with the additional requirement that $a = a'$ because forcing contracts compel the same action in every state.

FACT 6. The minimum punishment value for the setting of forcing contracts and ex post renegotiation is characterized as

$$P^{\mathrm{EPF}}(\theta, \theta') \equiv \min_{a \in A} \lambda(a, a, \theta, \theta').$$

Recall that proving Theorem 3 requires us to establish that $P^{\mathrm{EP}}(\theta, \theta') < P^{\mathrm{EPF}}(\theta, \theta')$ for some pair of states $\theta, \theta' \in \Theta$. Appendix C finishes the analysis by exploring how one can depart from the optimal forcing specification in a way that strictly reduces the value $\lambda(a, a', \theta, \theta')$.

6. Conclusion

In this paper, we report the analysis of contractual relationships for a large class of trade technologies. We highlight the usefulness of nonforcing contracts (in particular, the dual-option contract) and the key distinction between the *divided* and *unified* cases of investment and trade actions. We show that the payoff of the party with the trade action can be neutralized so that the other party claims the full benefit of the investment, gross of investment costs, implying that the efficient outcome is achieved in the divided case. Holdup remains a problem in the unified case, although the dual-option contract is optimal in a class of settings and it can sometimes induce the efficient outcome. We also provide general results on the relation between individual-action and public-action models of contractual relationships, showing that limiting attention to forcing contracts has significant implications for implementability and hence inefficiency.

Our results reinforce the message of Watson (2007) on the usefulness of modeling trade actions as individual, particularly in settings of cross investment. The results suggest revisiting some of the conclusions of public-action models in the existing literature. In particular, settings with cross investment are generally not as problematic as previous modeling exercises (Che and Hausch 1999, Edlin and Hermalin 2000, and others) find. Efficient outcomes can be achieved in the case of divided investment and trade actions. Our results show the importance, for applied work, of differentiating between the cases of divided and unified investment and trade actions. This distinction may be just as important as the distinction between own and cross investment (on which the literature has focused until now).

In our model, the trading opportunity is nondurable in that there is a single moment in time when trade can occur. One might wonder if the results differ substantially in settings with durable trading opportunities (where if trade does not occur at one time, then it can still be done at a later date). This issue is explored by Evans (2008) and Watson and Wignall (2009), both of whom examine individual-action models. Evans' (2008) elegant model is very general in terms of the available times at which the players can trade and renegotiate. He constructs equilibria in which, by having the players coordinate in different states on different equilibria in the infinite-horizon trade/negotiation game, the holdup problem is partly or completely alleviated. Evans' strongest result (in which the efficient outcome is reached) requires the ability of the players to commit to a joint financial hostage; that is, money is deposited with a third party until trade occurs, if ever.

Watson and Wignall (2009) examine a cross investment setting without the possibility of joint financial hostages, and their model is more modest in other dimensions. They show that the set of implementable post-investment payoff vectors in the setting of a durable trading opportunity is essentially the same as in the setting of a nondurable trading opportunity. This suggests that, in general, the results from the current paper carry over to the durability setting. Watson and Wignall also show that, in the divided case, there are nonstationary contracts that uniquely support the efficient outcome.

Our modeling exercise, combined with the recent literature, suggests some broad conclusions about the prospect of efficient investment and trade in contractual relationships. First, the holdup problem is not necessarily severe, and efficient outcomes can often be achieved. Durability of the trading opportunity does not worsen the holdup problem and may soften it in some cases, but it depends on the investment and trade technologies. Inefficiency may be unavoidable in the following problematic cases.

- When there is cross investment and unified investment and trade actions, as identified herein.
- When trade involves "complexity/ambivalence" as described by Segal (1999), Hart and Moore (1999), and Reiche (2006).
- When multiple parties make cross/cooperative investments.
- When the investment conveys a significant direct benefit (not requiring trade) on the noninvesting party, in addition to any benefit contingent on trade.

On the last point, Ellman's (2006) model provides intuition in terms of the notion of specificity. Settings in which multiple parties make cross investments are similar in nature to settings of team production (studied by Holmstrom 1982 and others).

In each of the cases above, the holdup problem would be reduced if the parties had some way to create joint financial hostages, as explored by Evans (2008), Boeckem and Schiller (2008), and Baliga and Sjöström (2009). Bull (2009) provides a cautionary note on the inability of such financial arrangements to withstand side contracting.

Regarding extensions of our analysis here, future research would be useful on different classes of trade technologies, in particular those in which both parties take trade actions (either simultaneously or sequentially). For instance, consider the setting with trade action profile $a = (a_1, a_2)$, where $a_1 \in [0, \overline{a}]$ is the verifiable quantity of an intermediate good that player 1 produces and delivers to player 2, and $a_2 \in \{\text{accept}, \text{reject}\}$ is player 2's verifiable choice of whether to accept or reject delivery. For simplicity, suppose that the players choose their trade actions simultaneously; the case of sequential choices works out similarly. Suppose that, for every state, "accept" is part of an ex post efficient trade-action profile.[19] Then a version of our results can be proved by utilizing contracts that force player 2 to accept delivery and are dual options with respect to a_1.

More precisely, if $(0, \text{accept})$ satisfies Assumption 1, then the conclusions of Theorem 1 and Proposition 1 go through, so an efficient outcome can be achieved in the divided case. Suppose further that, fixing $a_2 = \text{accept}$ and considering u as a function of a_1, Assumptions 2–4 hold. Then the conclusion of Theorem 2 is valid and the conclusions of Propositions 2 and 3 hold within the class of contracts that force player 2 to accept delivery, and are otherwise arbitrary (nonforcing and message-based).

It is not always possible to make player 1 the residual claimant by building a dual-option contract on player 2's acceptance/rejection choice, because there is typically not a single quantity a_1 that figures in the ex post efficient outcome in every state. Thus, we generally are not able to apply the argument for Theorem 1 to make player 1 the residual claimant. However, in some cases, a nonforcing specification for a_2 yields an improvement on the contracts that force acceptance of delivery.

[19] For instance, if positive trade is inefficient in some state, then $(0, \text{accept})$ is an ex post efficient outcome where player 1 delivers nothing and player 2 accepts.

There are technologies of trade for which it is possible to make either player the residual claimant with respect to the investment choice. Consider a setting with trade-action profile $a = (a_1, a_2)$ and, for any player i let j denote the other player. Here are assumptions that imply that player i can be made the residual claimant using a dual-option contract: Assume that there is a trade action $\hat{a}_i$ for player i and a trade action a_j^0 for player j such that, for every state θ, (i) $\hat{a}_i$ is an efficient action with an appropriate choice of a_j and (ii) $u_i(a_j^0, \hat{a}_i, \theta) = u_j(a_j^0, \hat{a}_i, \theta) = 0$. In words, the first condition says there is a trade action that player i can be forced to take, such that efficiency can be achieved in each state for some selection of player j's trade action. This condition is trivially satisfied in the setting of one-sided trade actions that we study. The second condition is Assumption 1 for j's trade action. Under these assumptions, player i can be made the residual claimant using the arguments behind Theorem 1. Clearly, it is possible for the assumptions to hold for both $i = 1$ and $i = 2$, although this seems to be a rather special case.

We have not begun an analysis of settings with more complicated multilateral trade actions, but we expect it to be a fruitful line of future research. Evans' (2008) model has a dynamic version of the trade technology described above, where one player makes a delivery choice and the other chooses whether to accept or reject delivery. It would also be interesting to look at a wide range of settings with partially verifiable trade actions. For example, a court may observe whether a particular trade was made but have trouble identifying which party disrupted trade (in the event that trade did not occur).[20]

Finally, recall that in the modeling exercise here, we assume that each party's productive actions are exogenously given. However, in some settings it may be possible to arbitrarily assign a particular task (such as delivering an object from one place to another) to an individual player. Our model indicates that the parties would have preferences over task assignment. Thus, it would be useful to determine whether physical trade actions are assignable in some real settings and to develop a model of optimal assignment. One might imagine a theory of firm boundaries that is based on the optimal assignment of different types of tasks over time.

Appendix A: Proof of Theorem 1

This appendix provides a proof of the first theorem. For any fixed k, consider the following contract. In the message phase (date 4), player 2 must declare the state. Let $\hat{\theta}$ denote player 2's announcement. If player 1 subsequently selects action $a^*(\hat{\theta})$, then the enforcer compels a transfer of $\hat{t} = (k - u_1(a^*(\hat{\theta}), \hat{\theta}), u_1(a^*(\hat{\theta}), \hat{\theta}) - k)$. If player 1 selects action a^0, then the transfer is $\underline{t} = (k, -k)$. If player 1 chooses any other trade action, then the enforcer compels transfer $(-\tau, \tau)$, where τ is set large enough so that player 1 is

[20] Hart and Moore's (1988) model has this feature. It is straightforward to incorporate partial verifiability into the modeling framework developed here. One can represent the external enforcer's information about the trading game as a partition of the space of action profiles. One can then simply assume that the contracted transfers y must be measurable with respect to this partition. Note that MacLeod and Malcomson (1993) and De Fraja (1999) examine settings with partially verifiable trade actions (along the lines of Hart and Moore 1988), although they make assumptions about the renegotiation protocol and the timing of outside options that weaken the affect of renegotiation compared to the rest of the literature.

forced to choose between $a^*(\hat{\theta})$ and a^0. That is, regardless of $\hat{\theta}$, in no state does player 1 have the incentive to choose $a \notin \{a^*(\hat{\theta}), a^0\}$.

Suppose that date 6 is reached without renegotiation and that the state is θ. Note that, by Assumption 1, player 1 gets a payoff of k if he chooses a^0. Alternatively, his payoff is

$$u_1(a^*(\hat{\theta}), \theta) + k - u_1(a^*(\hat{\theta}), \hat{\theta})$$

if he chooses $a^*(\hat{\theta})$. Thus, it is rational for player 1 to choose $a^*(\hat{\theta})$ if $u_1(a^*(\hat{\theta}), \theta) \geq u_1(a^*(\hat{\theta}), \hat{\theta})$ and to select a^0 otherwise, which we suppose is how player 1 behaves.

Consider next how player 2's payoff from date 4 depends on $\hat{\theta}$. Let θ be the actual state and divide the analysis into three cases. First, if player 2 declares $\hat{\theta} = \theta$, then, under the original contract, player 1 chooses $a^*(\hat{\theta})$ at date 6 and there is nothing to be jointly gained by renegotiating at date 5. In this case, the payoffs from date 4 are k for player 1 and

$$u_1(a^*(\theta), \theta) + u_2(a^*(\theta), \theta) - k = \gamma(\theta) - k$$

for player 2.

Second, if player 2 instead were to declare the state to be some $\hat{\theta} \neq \theta$ such that $u_1(a^*(\hat{\theta}), \theta) < u_1(a^*(\hat{\theta}), \hat{\theta})$, then the players anticipate that player 1 will select a^0 at date 6 under the original contract. Incorporating the impact of renegotiation at date 5, player 1's payoff from date 4 then is $k + \pi_1 R(a^0, \theta)$, where $R(a^0, \theta)$ is the renegotiation surplus in state θ if, without renegotiation, the players anticipate that a^0 will be the chosen trade action. Since $R(a^0, \theta) \geq 0$, player 1's payoff from date 4 weakly exceeds k and we conclude that player 2's payoff is weakly less than $\gamma(\theta) - k$.

Finally, suppose that player 2 declares the state to be $\hat{\theta} \neq \theta$ such that $u_1(a^*(\hat{\theta}), \theta) > u_1(a^*(\hat{\theta}), \hat{\theta})$. In this case, the players anticipate that player 1 will select $a^*(\hat{\theta})$ at date 6 under the original contract. Incorporating renegotiation at date 5, player 1's payoff from date 4 then is

$$u_1(a^*(\hat{\theta}), \theta) + k - u_1(a^*(\hat{\theta}), \hat{\theta}) + \pi_1 R(a^*(\hat{\theta}), \theta),$$

where $R(a^*(\hat{\theta}), \theta)$ is the renegotiation surplus in state θ if, without renegotiation, the players anticipate that $a^*(\hat{\theta})$ will be the chosen trade action. The first and third terms sum to weakly more than zero, so the entire expression weakly exceeds k. This implies that player 2's payoff is weakly less than $\gamma(\theta) - k$.

The foregoing analysis demonstrates that player 2 optimally tells the truth at date 4; that is, she declares $\hat{\theta} = \theta$. The payoffs from date 3 are thus k for player 1 and $\gamma(\theta) - k$ for player 2, which means that the contract implements the desired value function. □

Appendix B: Additional analysis and proofs for Section 3

In this appendix, we first provide necessary and sufficient conditions for making player 1 the residual claimant. We follow this with a note on when the conditions fail. We then provide a characterization of the punishment values defined in Sections 4 and 5. We use this characterization to examine the relation between the "inside conditions" and the "outside conditions" discussed in Section 3, which shows the difficulties of determining the second-best contract in the general unified case. Analysis of the inside and outside conditions for implementability motivates Theorem 2 and Propositions 2 and 3. This appendix concludes with their proofs.

Making player 1 the residual claimant

To make player 1 the residual claimant, we need to implement a value function v that satisfies, for some constant k, $v_2(\theta) = k$ and $v_1(\theta) = \gamma(\theta) - k$ for all $\theta \in \Theta$. Consider two states θ and θ', and order them so that $\theta > \theta'$. The conditions for implementation associated with these two states (for (θ, θ') and (θ', θ)) are

$$v_1(\theta) + v_2(\theta') \geq P^{\text{EP}}(\theta, \theta') \tag{8}$$

and

$$v_1(\theta') + v_2(\theta) \geq P^{\text{EP}}(\theta', \theta). \tag{9}$$

Using Fact 5 from Section 5, these conditions are equivalent to the existence of trade actions a, a', b, and b' such that $(a, a') \in E(\theta, \theta')$, $(b', b) \in E(\theta', \theta)$,

$$v_1(\theta) + v_2(\theta') \geq \lambda(a, a', \theta, \theta')$$

and

$$v_1(\theta') + v_2(\theta) \geq \lambda(b', b, \theta', \theta).$$

Substituting for v_1 and v_2 using the identities $v_2(\theta) = k$ and $v_1(\theta) = \gamma(\theta) - k$, these two inequalities become

$$\lambda(a, a', \theta, \theta') \leq \gamma(\theta) \tag{10}$$

and

$$\lambda(b', b, \theta', \theta) \leq \gamma(\theta'). \tag{11}$$

The following lemma summarizes.

Lemma 1. *Consider any contractual relationship that satisfies Assumptions 2(a) and 3. Let k be any real number and define value function v by $v_2(\theta) = k$ and $v_1(\theta) = \gamma(\theta) - k$ for all $\theta \in \Theta$. Then $v \in V^{\text{EP}}$ if and only if, for all pairs of states θ, θ' with $\theta > \theta'$, there are trade actions a, a', b, and b' such that $(a, a') \in E(\theta, \theta')$, $(b', b) \in E(\theta', \theta)$, and inequalities (10) and (11) hold.*

One can use these conditions to establish whether efficient investment can be obtained in specific examples with unified investment and trade actions, but sufficient conditions are stronger than are the assumptions we make in this paper.

For an illustration of cases where the conditions of Lemma 1 fail, suppose that the strict version of Assumption 3 is satisfied, meaning u_1 is strictly supermodular. Further suppose that Assumption 2 holds. Also suppose that U is strictly increasing in θ and that $U(\overline{a}, \overline{\theta}) > \gamma(\underline{\theta})$. That is, the joint value of the highest trade action in the highest state exceeds the maximal joint value in the lowest state (gross of investment cost).

Using (7), $U = u_1 + u_2$, and some algebra, we can rewrite inequality (11) as

$$\pi_1[U(b,\theta) - U(b',\theta')] \leq \pi_2[\gamma(\theta') - \gamma(\theta)] - [u_1(b,\theta') - u_1(b,\theta)].$$

Examining the case of $\theta = \overline{\theta}$ and $\theta' = \underline{\theta}$, this becomes

$$\pi_1[U(b,\overline{\theta}) - U(b',\underline{\theta})] \leq \pi_2[\gamma(\underline{\theta}) - \gamma(\overline{\theta})] - [u_1(b,\underline{\theta}) - u_1(b,\overline{\theta})]. \tag{12}$$

Because u_1 is strictly supermodular, $b \geq b'$ is required. From Assumption 2(b), that $U(\overline{a}, \overline{\theta}) > \gamma(\underline{\theta})$, and that U is strictly increasing in θ, we conclude that the left side of inequality (12) is strictly positive and bounded away from zero.[21] We also have that the first bracketed term on the right side is strictly negative.

Thus, if $|u_1(b,\underline{\theta}) - u_1(b,\overline{\theta})|$ is small relative to $\pi_2|\gamma(\underline{\theta}) - \gamma(\overline{\theta})|$, then inequality (12) fails to hold and there is no way to implement value functions that make player 2's payoff constant in the state. In other words, in the case of unified investment and trade actions, the first-best level of investment generally cannot be induced.

Optimal punishments and investment incentives

To construct the second-best contract in general, we first need to determine how to optimally punish deviations from truth-telling in the message phase (date 4). That is, we must calculate the punishment values. Although there are two implementation conditions for each pair of states, as shown above, we can focus on the one that bounds the rise in v_1 in the state. This condition corresponds to expression (11). Another way to look at this condition is to start with inequality (9), substitute for $v_2(\theta) = \gamma(\theta) - v_1(\theta)$, and rearrange terms to obtain

$$v_1(\theta) - v_1(\theta') \leq \gamma(\theta) - P^{\text{EP}}(\theta', \theta) \tag{13}$$

for $\theta > \theta'$. So lowering (improving) the punishment value for states (θ', θ) relaxes the constraint on how v_1 may rise with the state. Incidentally, from inequality (8), we get the corresponding bound for player 2:

$$v_2(\theta) - v_2(\theta') \leq \gamma(\theta) - P^{\text{EP}}(\theta, \theta').$$

These two inequalities are usefully employed to determine the limits of contracting in applications. For instance, we used the forcing-contract version (with P^{EPF}) to verify that the null contract is the optimal forcing contract in our example.

Our next step is to calculate $P^{\text{EP}}(\theta', \theta)$, the minimal punishment value, for any pair of states (θ', θ) with $\theta' < \theta$. This turns out to be straightforward under our assumptions. An alternative to Assumption 4(a) follows.

[21]To see this, consider two cases. If $U(\underline{a}, \overline{\theta}) \geq U(\overline{a}, \overline{\theta})$, because U is strictly quasiconcave in a, every point on the graph of $U(\cdot, \overline{\theta})$ is above every point on the graph of $U(\cdot, \underline{\theta})$ and so the result is immediate. If $U(\underline{a}, \overline{\theta}) < U(\overline{a}, \overline{\theta})$, U strictly increasing in θ implies that the result holds over the range $[\underline{a}, a^*(\overline{\theta})]$. Over the range $[a^*(\overline{\theta}), \overline{a}]$, the problem reduces to the first case.

Assumption 4. *(a′) For all* $\theta \in (\underline{\theta}, \overline{\theta}]$, $U(\underline{a}, \theta) \geq U(\overline{a}, \theta)$.

Lemma 2. *Under Assumptions 2, 3, and either Assumption 4(a) or 4(a′), for any pair of states* (θ', θ) *with* $\theta' < \theta$*, the optimal punishment involves inducing player 1 to select* $a^*(\theta')$ *in state* θ' *and* $\overline{a}$ *in state* θ*. That is,* $P^{\mathrm{EP}}(\theta', \theta) = \lambda(a^*(\theta'), \overline{a}, \theta', \theta)$.

Proof. From (7), the punishment value for (θ', θ), $\lambda(a', a, \theta', \theta)$, is given by

$$u_1(a, \theta') + \pi_1 R(a', \theta') + u_2(a, \theta) + \pi_2 R(a, \theta),$$

which can be rewritten as

$$u_1(a, \theta') + \pi_1 U(a^*(\theta'), \theta') - \pi_1 U(a', \theta') + \pi_1 u_2(a, \theta) + \pi_2 U(a^*(\theta), \theta) - \pi_2 u_1(a, \theta).$$

The optimal punishment value is obtained by choosing a and a' to minimize this objective function under the constraint that $a \geq a'$ (because of supermodularity of u_1). Ignoring the constant terms that do not contain a or a', and substituting $\pi_2 = (1 - \pi_1)$ and $\pi_1 u_1(a, \theta) + \pi_1 u_2(a, \theta) = \pi_1 U(a, \theta)$, the objective function becomes

$$-[u_1(a, \theta) - u_1(a, \theta')] + \pi_1 U(a, \theta) - \pi_1 U(a', \theta'). \tag{14}$$

The number a' affects only the last term; to minimize it (that is, maximize $U(a', \theta')$) without consideration of the constraint $a \geq a'$, it is optimal to set $a' = a^*(\theta')$. From supermodularity of u_1, the negative bracketed term is minimized by choosing $a = \overline{a}$. By Assumption 4(a′), the final term is also minimized at $\overline{a}$. Thus, $\lambda(a', a, \theta', \theta)$ attains its lowest value when $a = \overline{a}$ and $a' = a^*(\theta')$.

To see that the same result holds with Assumption 4(a) in place of Assumption 4(a′), observe that if the minimizing values a and a' satisfy $a > a'$ then it must be that $a = \overline{a}$ and $a' = a^*(\theta')$. That $a' = a^*(\theta')$ is an implication of strict quasiconcavity of $U(\cdot, \theta')$, for if $a' < a$ and $a' \neq a^*(\theta')$, then it must be that $a^*(\theta') > a$, but then raising a' to a strictly increases $U(a', \theta')$. The conclusion that $a = \overline{a}$ follows from strict quasiconcavity of $U(\cdot, \theta)$ and from supermodularity of u_1, which imply that it is not optimal to set $a \in (a', \overline{a})$.

So we know that either it is optimal to have $a = \overline{a}$ and $a' = a^*(\theta')$, or $a = a'$ is optimal. In the latter case, Assumption 4(a) implies that $a = a' = \overline{a}$ is best. This is apparent by rearranging terms to show that, by substituting $a = a'$ into expression (14), the expression becomes

$$\pi_1[u_2(a, \theta) - u_2(a, \theta')] - \pi_2[u_1(a, \theta) - u_1(a, \theta')],$$

which is decreasing in a. This yields a contradiction because we can lower a' to $a^*(\theta')$ to strictly decrease the objective function. □

A note about inside and outside constraints on value functions

Lemma 2 allows us to easily calculate the lowest possible punishment values for any unilateral deviation from truth-telling, and with this in hand we can begin to evaluate how the conditions for implementability come together to constrain the value function. For the unified case, we want to know whether we can implement a value function so that v_1 increases at the same rate as does γ. One way to get at this is to examine constraints on $v_1(\theta) - v_1(\theta')$ at the margin where θ and θ' are very close, and then chain together these inside conditions to characterize the optimal implementable value function.

Unfortunately, there are also outside conditions to examine; they give constraints on $v_1(\theta) - v_1(\theta')$ for θ and θ' that are far apart. We demonstrate that the outside conditions are typically tighter than the sum of the inside conditions, so a triangle inequality fails. Thus, one cannot rely on marginal analysis to calculate bounds on implementable value functions (a cautionary note relative to Segal and Whinston 2002).

Consider any three states satisfying $\theta^{\mathrm{L}} < \theta^{\mathrm{M}} < \theta^{\mathrm{H}}$, such that the optimal trade action in state θ^{M} is interior so that $a^*(\theta^{\mathrm{M}}) \in (\underline{a}, \overline{a})$. Also assume that the assumptions for Lemma 2 hold. Using inequality (13), we have three necessary conditions for implementation:

$$\begin{aligned} v_1(\theta^{\mathrm{M}}) - v_1(\theta^{\mathrm{L}}) &\le \gamma(\theta^{\mathrm{M}}) - P^{\mathrm{EP}}(\theta^{\mathrm{L}}, \theta^{\mathrm{M}}) \\ v_1(\theta^{\mathrm{H}}) - v_1(\theta^{\mathrm{M}}) &\le \gamma(\theta^{\mathrm{H}}) - P^{\mathrm{EP}}(\theta^{\mathrm{M}}, \theta^{\mathrm{H}}) \\ v_1(\theta^{\mathrm{H}}) - v_1(\theta^{\mathrm{L}}) &\le \gamma(\theta^{\mathrm{H}}) - P^{\mathrm{EP}}(\theta^{\mathrm{L}}, \theta^{\mathrm{H}}). \end{aligned} \tag{15}$$

Summing the first two yields

$$v_1(\theta^{\mathrm{H}}) - v_1(\theta^{\mathrm{L}}) \le \gamma(\theta^{\mathrm{H}}) - P^{\mathrm{EP}}(\theta^{\mathrm{M}}, \theta^{\mathrm{H}}) + \gamma(\theta^{\mathrm{M}}) - P^{\mathrm{EP}}(\theta^{\mathrm{L}}, \theta^{\mathrm{M}}). \tag{16}$$

We want to know whether (16) is a weakly tighter bound than is (15), which would mean that the inside conditions imply the outside conditions and allow implementability to be characterized by marginal analysis. So we must establish whether the following triangle inequality holds:

$$P^{\mathrm{EP}}(\theta^{\mathrm{L}}, \theta^{\mathrm{H}}) + \gamma(\theta^{\mathrm{M}}) \le P^{\mathrm{EP}}(\theta^{\mathrm{L}}, \theta^{\mathrm{M}}) + P^{\mathrm{EP}}(\theta^{\mathrm{M}}, \theta^{\mathrm{H}}).$$

Expanding terms using expression (7) and Lemma 2, we get

$$\begin{aligned} &u_1(\overline{a}, \theta^{\mathrm{L}}) + u_2(\overline{a}, \theta^{\mathrm{H}}) + \pi_2 R(\overline{a}, \theta^{\mathrm{H}}) + \gamma(\theta^{\mathrm{M}}) \\ &\qquad \le u_1(\overline{a}, \theta^{\mathrm{L}}) + u_2(\overline{a}, \theta^{\mathrm{M}}) + \pi_2 R(\overline{a}, \theta^{\mathrm{M}}) + u_1(\overline{a}, \theta^{\mathrm{M}}) + u_2(\overline{a}, \theta^{\mathrm{H}}) + \pi_2 R(\overline{a}, \theta^{\mathrm{H}}), \end{aligned}$$

which simplifies to

$$\gamma(\theta^{\mathrm{M}}) \le u_2(\overline{a}, \theta^{\mathrm{M}}) + \pi_2 R(\overline{a}, \theta^{\mathrm{M}}) + u_1(\overline{a}, \theta^{\mathrm{M}}).$$

This is equivalent to

$$\gamma(\theta^{\mathrm{M}}) \le U(\overline{a}, \theta^{\mathrm{M}}) + \pi_2 \gamma(\theta^{\mathrm{M}}) - \pi_2 U(\overline{a}, \theta^{\mathrm{M}}),$$

which simplifies to $\gamma(\theta^{\mathrm{M}}) \le U(\overline{a}, \theta^{\mathrm{M}})$. This inequality, coupled with Assumption 2(b), requires that $a^*(\theta^{\mathrm{M}}) = \overline{a}$, which contradicts what we assumed earlier.

Proof of Theorem 2

Consider the following contract: In the message phase (date 4), player 2 must declare the state. Let $\hat{\theta}$ denote player 2's announcement. If $\hat{\theta} \geq \beta$, then player 1 is forced to select $\overline{a}$ at date 6. Otherwise, player 1 is forced to choose between $a^*(\hat{\theta})$ and $\overline{a}$. In this case, if player 1 selects action $a^*(\hat{\theta})$, then the enforcer compels a transfer of

$$\hat{t} = \big(u_1(\overline{a}, \hat{\theta}) - u_1(a^*(\hat{\theta}), \hat{\theta}), u_1(a^*(\hat{\theta}), \hat{\theta}) - u_1(\overline{a}, \hat{\theta})\big), \tag{17}$$

and if player 1 selects action $\overline{a}$, then the transfer is $\overline{t} = (0, 0)$. The forcing arrangement is achieved by specifying a transfer of $(-\tau, \tau)$ if player 1 picks any other trade action, where τ is set large enough to keep him from doing so.

We show that this contract implements the value function v^β defined in the text. First note that, in any state θ, if at date 4 player 2 declares the state to be $\hat{\theta} \in [\underline{\theta}, \beta)$ and the players do not renegotiate at date 5, then player 1 obtains at least $u_1(\overline{a}, \theta)$ because he has the option of choosing $\overline{a}$ with no transfer. Further, if player 1 selects $a^*(\hat{\theta})$, then (from expression (17)) he gets

$$u_1(a^*(\hat{\theta}), \theta) + u_1(\overline{a}, \hat{\theta}) - u_1(a^*(\hat{\theta}), \hat{\theta}),$$

whereas he gets $u_1(\overline{a}, \theta)$ by choosing $\overline{a}$. The latter payoff weakly exceeds the former if and only if

$$u_1(\overline{a}, \theta) - u_1(a^*(\hat{\theta}), \theta) \geq u_1(\overline{a}, \hat{\theta}) - u_1(a^*(\hat{\theta}), \hat{\theta}).$$

From the supermodularity of u_1 and given that $\overline{a} \geq a^*(\hat{\theta})$, we know that it is rational for player 1 to choose $\overline{a}$ in the case of $\hat{\theta} < \theta$ and it is rational for player 1 to choose $a^*(\hat{\theta})$ in the case of $\hat{\theta} > \theta$. Player 1 is indifferent if $\hat{\theta} = \theta$.

We can therefore prescribe the following behavior, for any state θ.

- If player 2 declares $\hat{\theta} \in [\theta, \beta)$, then, absent renegotiation, player 1 chooses $a^*(\hat{\theta})$ at date 6.
- For any other message (either $\hat{\theta} \geq \beta$ or $\hat{\theta} < \theta$), absent renegotiation, player 1 chooses $\overline{a}$.

It is clear that, given player 1's behavior just specified, it is optimal for player 2 to report truthfully at date 4. For instance, in a state $\theta < \beta$, if player 2 reports honestly, then there is no renegotiation and she gets $\gamma(\theta) - u_1(\overline{a}, \theta)$. If she reports a different state, then player 1 is expected to take an ex post inefficient trade action that gives him at least $u_1(\overline{a}, \theta)$, so player 2 fares less well.

With the specified behavior for the players, in any state $\theta < \beta$ there is no renegotiation and player 1 obtains the payoff $u_1(\overline{a}, \theta)$. In any state $\theta \geq \beta$, the players renegotiate away from the anticipated action of $\overline{a}$ and player 1 gets $u_1(\overline{a}, \theta) + \pi_1 R(\overline{a}, \theta)$. Thus, value function v^β is implemented. □

Proof of Proposition 2

Consider any two states θ' and θ with $\theta' < \theta$. Using Lemma 2 for the pairing (θ', θ), which shows that the minimum punishment value for such a pair involves inducing player 1 to select $a^*(\theta')$ in state θ' and $\overline{a}$ in state θ, we have that

$$P^{\mathrm{EP}}(\theta', \theta) = u_1(\overline{a}, \theta') + u_2(\overline{a}, \theta) + \pi_2 R(\overline{a}, \theta).$$

We use this equality to substitute for P^{EP} in the necessary condition (13). Rearranging terms yields the following upper bound on the value difference between the two states for player 1:

$$v_1(\theta) - v_1(\theta') \leq \pi_1 R(\overline{a}, \theta) + u_1(\overline{a}, \theta) - u_1(\overline{a}, \theta'). \tag{18}$$

Consider any contract that supports the best achievable investment level θ^{B} and let v^{B} be the implemented value function. By definition of θ^{B}, we know that θ^{B} solves player 1's investment problem of maximizing $v_1^{\mathrm{B}}(\theta) - \theta$.

Define $\beta \equiv \theta^{\mathrm{B}}$ and consider the value function v^β defined in Theorem 2. To see that v^β supports θ^{B} (that is, θ^{B} maximizes $v_1^\beta(\theta) - \theta$), first observe that for any state $\theta < \theta^{\mathrm{B}}$, we have

$$v_1^\beta(\theta^{\mathrm{B}}) - v_1^\beta(\theta) = u_1(\overline{a}, \theta^{\mathrm{B}}) + \pi_1 R(\overline{a}, \theta^{\mathrm{B}}) - u_1(\overline{a}, \theta),$$

so v^β meets the upper bound on player 1's payoff difference between θ and θ^{B}, as identified in inequality (18). Since v^{B} also must satisfy the bound (18), we conclude that $v_1^\beta(\theta^{\mathrm{B}}) - v_1^\beta(\theta) \geq v_1^{\mathrm{B}}(\theta^{\mathrm{B}}) - v_1^{\mathrm{B}}(\theta)$ for all $\theta < \theta^{\mathrm{B}}$. This implies that the maximizer of $v_1^\beta(\theta) - \theta$ must be no less than θ^{B}.

The final step is to consider the implications of θ^{B} not maximizing $v_1^\beta(\theta) - \theta$. In this case, let $\tilde{\theta} > \theta^{\mathrm{B}}$ denote an investment that player 1 strictly prefers. We then have $v_1^\beta(\tilde{\theta}) - \tilde{\theta} > v_1^\beta(\theta^{\mathrm{B}}) - \theta^{\mathrm{B}}$. Plugging in the implemented values of v_1^β, this is equivalent to

$$u_1(\overline{a}, \tilde{\theta}) + \pi_1 R(\overline{a}, \tilde{\theta}) - \tilde{\theta} > u_1(\overline{a}, \theta^{\mathrm{B}}) + \pi_1 R(\overline{a}, \theta^{\mathrm{B}}) - \theta^{\mathrm{B}}.$$

Rearranging terms, we see that this is equivalent to

$$\pi_1 \gamma(\tilde{\theta}) - \tilde{\theta} - [\pi_1 \gamma(\theta^{\mathrm{B}}) - \theta^{\mathrm{B}}] > \pi_2 U(\overline{a}, \theta^{\mathrm{B}}) - \pi_2 U(\overline{a}, \tilde{\theta}) + u_2(\overline{a}, \tilde{\theta}) - u_2(\overline{a}, \theta^{\mathrm{B}}).$$

It is not difficult to verify that Assumption 4(b) implies that the expression on the right side is weakly positive and thus the expression on the left side is strictly positive. This further implies that $\gamma(\tilde{\theta}) > \gamma(\theta^{\mathrm{B}})$, which means that

$$\gamma(\tilde{\theta}) - \tilde{\theta} - [\gamma(\theta^{\mathrm{B}}) - \theta^{\mathrm{B}}] > 0,$$

contradicting that θ^{B} is the best achievable investment level. Thus, we know that θ^{B} maximizes $v_1^\beta(\theta) - \theta$. □

Proof of Proposition 3

If θ^* is supported, then Proposition 2 implies that it is supported by the value function v^{θ^*} from Theorem 2. We then know that θ^* maximizes $v_1^{\theta^*}(\theta) - \theta$ by choice of θ, which implies that inequality (2) holds for all $\theta < \theta^*$. Thus, the condition of the proposition is necessary. Sufficiency requires not only that inequality (2) hold for all $\theta < \theta^*$, but also that player 1 prefer not to invest $\tilde{\theta} > \theta^*$ when v^{θ^*} is the implemented value function. This follows from the argument in the final paragraph of the proof of Proposition 2, replacing θ^{B} with θ^* and "best achievable" with "efficient." □

Appendix C: Proof of Theorem 3

In this appendix, we complete the proof of Theorem 3. We start with the comparison of V^{EPF} and V^{EP} and then provide the analysis for the comparison of V^{EP} and V^{I}.

Completion of the proof that $V^{\mathrm{EPF}} \neq V^{\mathrm{EP}}$

We pick up from the analysis at the end of Section 5. Consider a pair of states θ^1 and θ^2 that satisfies Assumption 5. That is, we have $\theta^1 > \theta^2$ and either $U(\underline{a}, \theta^2) < U(\overline{a}, \theta^2)$ or $U(\underline{a}, \theta^1) > U(\overline{a}, \theta^1)$. Let b^1 denote a solution to the forcing-contract problem

$$\min_{a \in A} \lambda(a, a, \theta^1, \theta^2)$$

and let b^2 denote a solution to the forcing-contract problem

$$\min_{a \in A} \lambda(a, a, \theta^2, \theta^1).$$

It is easy to show that $b^1 \geq a^*(\theta^1) > \underline{a}$ and $b^2 \leq a^*(\theta^2) < \overline{a}$ follow from Assumptions 2(b) and 5. We use these facts below. We demonstrate that either $P^{\mathrm{EP}}(\theta^1, \theta^2) < P^{\mathrm{EPF}}(\theta^1, \theta^2)$ or $P^{\mathrm{EP}}(\theta^2, \theta^1) < P^{\mathrm{EPF}}(\theta^2, \theta^1)$ or both, which implies that $V^{\mathrm{EPF}} \neq V^{\mathrm{EP}}$.

Let us evaluate the minimum punishment value that corresponds to the ordered pair of states (θ^1, θ^2). Specifically, compare the optimal forcing-contract punishment (forcing player 1 to select b^1 in both states) with a nonforcing specification in which player 1 is induced to select b^1 in state θ^1 and $\underline{a}$ in state θ^2. This is a valid nonforcing contractual specification because, by Fact 2, $\theta^1 > \theta^2$ and $b^1 > \underline{a}$ imply $(b^1, \underline{a}) \in E(\theta^1, \theta^2)$.

If $V^{\mathrm{EP}} = V^{\mathrm{EPF}}$, then it must be that $\lambda(b^1, b^1, \theta^1, \theta^2) \leq \lambda(b^1, \underline{a}, \theta^1, \theta^2)$. Applying the definition of λ, this is

$$u_1(b^1, \theta^1) + \pi_1 R(b^1, \theta^1) + u_2(b^1, \theta^2) + \pi_2 R(b^1, \theta^2) \\ \leq u_1(\underline{a}, \theta^1) + \pi_1 R(b^1, \theta^1) + u_2(\underline{a}, \theta^2) + \pi_2 R(\underline{a}, \theta^2).$$

Canceling the second term on each side and using the definition of R, we arrive at

$$u_1(b^1, \theta^1) + u_2(b^1, \theta^2) - \pi_2 U(b^1, \theta^2) \leq u_1(\underline{a}, \theta^1) + u_2(\underline{a}, \theta^2) - \pi_2 U(\underline{a}, \theta^2).$$

Substituting $u_2(\cdot, \theta^2) = U(\cdot, \theta^2) - u_1(\cdot, \theta^2)$ on both sides, we have

$$u_1(b^1, \theta^1) + U(b^1, \theta^2) - u_1(b^1, \theta^2) - \pi_2 U(b^1, \theta^2) \\ \leq u_1(\underline{a}, \theta^1) + U(\underline{a}, \theta^2) - u_1(\underline{a}, \theta^2) - \pi_2 U(\underline{a}, \theta^2).$$

Finally, rearranging this expression a bit and using $\pi_1 + \pi_2 = 1$, we conclude that $\lambda(b^1, b^1, \theta^1, \theta^2) \leq \lambda(b^1, \underline{a}, \theta^1, \theta^2)$ is equivalent to

$$u_1(b^1, \theta^1) - u_1(\underline{a}, \theta^1) - [u_1(b^1, \theta^2) - u_1(\underline{a}, \theta^2)] \leq \pi_1[U(\underline{a}, \theta^2) - U(b^1, \theta^2)]. \tag{23}$$

Similarly, ordering states θ^1 and θ^2 in the opposite way, we compare the optimal forcing-contract punishment (forcing player 1 to select b^2 in both states) with a non-forcing specification in which player 1 is induced to select b^2 in state θ^2 and $\overline{a}$ in state θ^1. Note that $\theta^2 < \theta^1$ and $b^2 < \overline{a}$ imply $(b^2, \overline{a}) \in E(\theta^2, \theta^1)$. If $V^{\mathrm{EP}} = V^{\mathrm{EPF}}$, then it must be that $\lambda(b^2, b^2, \theta^2, \theta^1) \leq \lambda(b^2, \overline{a}, \theta^2, \theta^1)$, which similar algebraic manipulation reveals to be equivalent to

$$u_1(\overline{a}, \theta^1) - u_1(b^2, \theta^1) - [u_1(\overline{a}, \theta^2) - u_1(b^2, \theta^2)] \leq \pi_1[U(\overline{a}, \theta^1) - U(b^2, \theta^1)]. \tag{24}$$

The foregoing analysis shows that if $V^{\mathrm{EPF}} = V^{\mathrm{EP}}$, then expressions (23) and (24) hold. Assumption 3 then implies that the left sides of these inequalities are nonnegative, which implies

$$U(\underline{a}, \theta^2) \geq U(b^1, \theta^2) \quad \text{and} \quad U(\overline{a}, \theta^1) \geq U(b^2, \theta^1).$$

Using Assumption 2(b), and that $b^1 > \underline{a}$ and $b^2 < \overline{a}$, we obtain the following fact.

Fact 7. If $V^{\mathrm{EPF}} = V^{\mathrm{EP}}$, then $U(\underline{a}, \theta^2) \geq U(\overline{a}, \theta^2)$ and $U(\overline{a}, \theta^1) \geq U(\underline{a}, \theta^1)$.

Assumption 5 and the contrapositive of Fact 7 provide the contradiction that proves $V^{\mathrm{EPF}} \neq V^{\mathrm{EP}}$.

Proof that $V^{\mathrm{EP}} \neq V^{\mathrm{I}}$

We next prove the claim about the relation between V^{I} and V^{EP}. Since forcing contracts are sufficient to construct V^{I}, we can state the following fact.

Fact 8. The minimum punishment value in the setting of interim renegotiation is characterized as

$$P^{\mathrm{I}}(\theta, \theta') = \min_{a'' \in A} u_1(a'', \theta) + u_2(a'', \theta').$$

Remember that, by Result 2, $V^{\mathrm{I}} = V^{\mathrm{EP}}$ if and only if $P^{\mathrm{EP}}(\theta, \theta') = P^{\mathrm{I}}(\theta, \theta')$ for all $\theta, \theta' \in \Theta$. We can again compare the minimization problems to determine if this is the case.

Take θ^1 and θ^2 as satisfying Assumption 5. Consider any solution to the minimization problem that defines $P^{\text{EP}}(\theta^1, \theta^2)$ and denote it (b, b'). That is, (b, b') solves

$$\min_{(a,a')\in E(\theta^1,\theta^2)} u_1(a', \theta^1) + \pi_1 R(a, \theta^1) + u_2(a', \theta^2) + \pi_2 R(a', \theta^2).$$

Then $P^{\text{EP}}(\theta^1, \theta^2) = P^{\text{I}}(\theta^1, \theta^2)$ is equivalent to

$$u_1(b', \theta^1) + \pi_1 R(b, \theta^1) + u_2(b', \theta^2) + \pi_2 R(b', \theta^2) = \min_{a''\in A} u_1(a'', \theta^1) + u_2(a'', \theta^2).$$

Because $R(\cdot, \cdot) \geq 0$, we see that $P^{\text{EP}}(\theta^1, \theta^2) = P^{\text{I}}(\theta^1, \theta^2)$ only if b' solves the minimization problem on the right side of the above equation and also $R(b, \theta^1) = R(b', \theta^2) = 0$.

By Assumption 2(b), $R(b', \theta^2) = 0$ if and only if $b' = a^*(\theta^2)$. Combining this with the requirement that b' must minimize $u_1(\cdot, \theta^1) + u_2(\cdot, \theta^2)$, we derive that

$$u_1(a^*(\theta^2), \theta^1) + u_2(a^*(\theta^2), \theta^2) \leq u_1(a'', \theta^1) + u_2(a'', \theta^2)$$

for all a''. In particular, the following inequality must hold:

$$u_1(a^*(\theta^2), \theta^1) + u_2(a^*(\theta^2), \theta^2) \leq u_1(\underline{a}, \theta^1) + u_2(\underline{a}, \theta^2).$$

Using the identity $u_2 = U - u_1$ and rearranging terms, we see that this is equivalent to

$$\begin{aligned} u_1(a^*(\theta^2), \theta^1) - u_1(\underline{a}, \theta^1) - \big[u_1(a^*(\theta^2), \theta^2) - u_1(\underline{a}, \theta^2)\big] \\ \leq U(\underline{a}, \theta^2) - U(a^*(\theta^2), \theta^2). \end{aligned} \tag{25}$$

Similarly, ordering states θ^1 and θ^2 in the opposite way, it is necessary that $a^*(\theta^1)$ must solve $P^{\text{I}}(\theta^2, \theta^1)$ so that $P^{\text{EP}}(\theta^2, \theta^1) = P^{\text{I}}(\theta^2, \theta^1)$. In particular, we must have

$$u_1(a^*(\theta^1), \theta^2) + u_2(a^*(\theta^1), \theta^1) \leq u_1(\overline{a}, \theta^2) + u_2(\overline{a}, \theta^1).$$

This inequality is equivalent to

$$\begin{aligned} u_1(\overline{a}, \theta^1) - u_1(a^*(\theta^1), \theta^1) - \big[u_1(\overline{a}, \theta^2) - u_1(a^*(\theta^1), \theta^2)\big] \\ \leq U(\overline{a}, \theta^1) - U(a^*(\theta^1), \theta^1). \end{aligned} \tag{26}$$

By Assumption 3, the left sides of expressions (25) and (26) must be nonnegative, which implies both $U(\underline{a}, \theta^2) \geq U(a^*(\theta^2), \theta^2)$ and $U(\overline{a}, \theta^1) \geq U(a^*(\theta^1), \theta^1)$. From Assumption 2(b), we see that this is only possible if $\underline{a} = a^*(\theta^2)$ and $\overline{a} = a^*(\theta^1)$. If this is the case, Assumption 2(b) also implies that $U(\underline{a}, \theta^2) \geq U(\overline{a}, \theta^2)$ and $U(\overline{a}, \theta^1) \geq U(\underline{a}, \theta^1)$. Thus we obtain our last fact.

Fact 9. If $V^{\text{I}} = V^{\text{EP}}$, then $U(\underline{a}, \theta^2) \geq U(\overline{a}, \theta^2)$ and $U(\overline{a}, \theta^1) \geq U(\underline{a}, \theta^1)$.

The contrapositive of Fact 9 combined with Assumption 5 provides the contradiction that proves $V^{\text{I}} \neq V^{\text{EP}}$. □

References

Aghion, Philippe, Mathias Dewatripont, and Patrick Rey (1994), "Renegotiation design with unverifiable information." *Econometrica*, 62, 257–282. [290, 291, 294]

Baliga, Sandeep and Tomas Sjöström (2009), "Contracting with third parties." *American Economic Journal: Microeconomics*, 1, 75–100. [308]

Beaudry, Paul and Michel Poitevin (1995), "Contract renegotiation: A simple framework and implications for organization theory." *Canadian Journal of Economics*, 28, 302–335. [290]

Boeckem, Sabine and Ulf Schiller (2008), "Option contracts in supply chains." *Journal of Economics and Management Strategy*, 17, 219–245. [290, 308]

Bull, Jesse (2009), "Third-party budget breakers and side contracting in team production." Unpublished paper, Florida International University. [308]

Che, Yeon Koo and Donald B. Hausch (1999), "Cooperative investments and the value of contracting." *American Economic Review*, 89, 125–147. [285, 290, 291, 292, 307]

Che, Yeon Koo and József Sákovics (2008), "Hold-up problem." In *The New Palgrave Dictionary of Economics* (Lawrence E. Bloom and Steven N. Durlauf, eds.), Palgrave Macmillan, Hampshire, UK. [283]

Chung, Tai Yeong (1991), "Incomplete contracts, specific investments, and risk sharing." *Review of Economic Studies*, 58, 1031–1042. [291]

De Fraja, Gianni (1999), "After you sir. Hold-up, direct externalities, and sequential investment." *Games and Economic Behavior*, 26, 22–39. [309]

Demski, Joel S. and David E. M. Sappington (1991), "Resolving double moral hazard problems with buyout agreements." *RAND Journal of Economics*, 22, 232–240. [290, 292]

Edlin, Aaron S. and Benjamin E. Hermalin (2000), "Contract renegotiation and options in agency problems." *Journal of Law, Economics, and Organization*, 16, 395–423. [290, 307]

Edlin, Aaron S. and Stefan Reichelstein (1996), "Holdups, standard breach remedies, and optimal investment." *American Economic Review*, 86, 478–501. [290, 291]

Ellman, Matthew (2006), "Specificity revisited: The role of cross-investments." *Journal of Law, Economics, and Organization*, 22, 234–257. [290, 308]

Evans, Robert (2006), "Mechanism design with renegotiation and costly messages." Unpublished paper, University of Cambridge. [290]

Evans, Robert (2008), "Simple efficient contracts in complex environments." *Econometrica*, 76, 459–491. [290, 307, 308, 309]

Grossman, Sanford J. and Oliver D. Hart (1986), "The costs and benefits of ownership: A theory of vertical and lateral integration." *Journal of Political Economy*, 94, 691–719. [283]

Grout, Paul A. (1984), "Investment and wages in the absence of binding contracts: A Nash bargaining approach." *Econometrica*, 52, 449–460. [283]

Guriev, Sergei (2003), "Incomplete contracts with cross-investments." *Contributions in Theoretical Economics*, 3. [290]

Hart, Oliver D. and John H. Moore (1988), "Incomplete contracts and renegotiation." *Econometrica*, 56, 755–785. [283, 290, 292, 309]

Hart, Oliver D. and John H. Moore (1999), "Foundations of incomplete contracts." *Review of Economic Studies*, 66, 115–138. [290, 291, 308]

Holmstrom, Bengt (1982), "Moral hazard in teams." *Bell Journal of Economics*, 13, 324–340. [308]

Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian (1978), "Vertical integration, appropriable rents, and the competitive contracting process." *Journal of Law and Economics*, 21, 297–326. [283]

Lyon, Thomas P. and Eric Rasmusen (2004), "Buyer-option contracts restored: Renegotiation, inefficient threats, and the hold-up problem." *Journal of Law, Economics, and Organization*, 20, 148–169. [290]

MacLeod, W. Bentley and James M. Malcomson (1993), "Investments, holdup, and the form of market contracts." *American Economic Review*, 83, 811–837. [290, 292, 294, 309]

Maskin, Eric S. and John H. Moore (1999), "Implementation and renegotiation." *Review of Economic Studies*, 66, 39–56. [284, 290]

Maskin, Eric S. and Jean Tirole (1999), "Two remarks on the property-rights literature." *Review of Economic Studies*, 66, 139–149. [290]

Myerson, Roger B. (1982), "Optimal coordination mechanisms in generalized principal–agent problems." *Journal of Mathematical Economics*, 10, 67–81. [290]

Myerson, Roger B. (1991), *Game Theory*. Harvard University Press, Cambridge, Massachusetts. [290]

Nöldeke, Georg and Klaus M. Schmidt (1995), "Option contracts and renegotiation: A solution to the hold-up problem." *RAND Journal of Economics*, 26, 163–179. [290, 291, 292]

Nöldeke, Georg and Klaus M. Schmidt (1998), "Sequential investments and options to own." *RAND Journal of Economics*, 29, 633–653. [290]

Reiche, Sönje (2006), "Ambivalent investment and the hold-up problem." *Journal of the European Economic Association*, 4, 1148–1164. [291, 308]

Rogerson, William P. (1992), "Contractual solutions to the hold-up problem." *Review of Economic Studies*, 59, 777–793. [291]

Roider, Andreas (2004), "Asset ownership and contractibility of interaction." *RAND Journal of Economics*, 35, 787–802. [290]

Segal, Ilya (1999), "Complexity and renegotiation: A foundation for incomplete contracts." *Review of Economic Studies*, 66, 57–82. [290, 291, 308]

Segal, Ilya and Michael D. Whinston (2002), "The Mirrlees approach to mechanism design with renegotiation (with applications to hold-up and risk sharing)." *Econometrica*, 70, 1–45. [284, 290, 292, 314]

Stremitzer, Alexander (forthcoming), "Standard breach remedies, quality thresholds, and cooperative investments." *Journal of Law, Economics, and Organization*. [290]

Watson, Joel (2005), "Contract and mechanism design in settings with multi-period trade." Unpublished paper, University of California, San Diego. [303]

Watson, Joel (2006), "Contract and game theory: Basic concepts for settings with finite horizons." Unpublished paper, University of California, San Diego. [288]

Watson, Joel (2007), "Contract, mechanism design, and technological detail." *Econometrica*, 75, 55–81. [284, 285, 286, 289, 290, 292, 300, 302, 303, 307]

Watson, Joel and Chris Wignall (2009), "Hold-up and durable trading opportunities." Unpublished paper, University of California, San Diego. [307]

Williamson, Oliver E. (1975), *Markets and Hierarchies: Analysis and Antitrust Implications*. Free Press, New York. [283]

Williamson, Oliver E. (1979), "Transaction cost economics: The governance of contractual relations." *Journal of Law and Economics*, 22, 233–261. [283]

Submitted 2010-6-21. Final version accepted 2011-6-19. Available online 2011-6-19.

Chapter 1, in full, is a reprint of the material as it appears in *Theoretical Economics*, 7, 2012, Buzard, Kristy; Watson, Joel, 2012. The dissertation author was a principal author of this paper.

Chapter 2

Trade Agreements, lobbying and separation of powers

Abstract

This paper presents a model of international trade agreements in which the executive branches of each government negotiate agreements while the legislative branches, subject to political pressure, can disrupt them. Lobbying is in the style of Grossman and Helpman (1994) with a new feature: all actors face uncertainty arising from the complexity of the legislative process. I demonstrate that the lower the executives set trade agreement tariffs, the more effort lobbies put forth to induce the legislatures to disrupt the agreement. Thus trade agreements act as a domestic political commitment device: executives set relatively high tariffs to discourage lobbying and help the legislatures to withstand political pressure. This reconciles the result from tests of Grossman and Helpman's model that protection levels are high relative to contributions given estimates of governments' social-welfare weights. Moreover, this rich modeling of the political process reveals that tariffs, lobbying and the likelihood of trade disruptions vary systematically with how much uncertainty there is about the weight the median legislators place on the profits of politically-organized industries.

KEYWORDS: Trade Policy, International Agreements, Lobbying, Political Economy, Structure of Government

Empirical investigations of Grossman and Helpman's 1994 'Protection for Sale' model (henceforth 'PFS') broadly support the predictions of the theory qualitatively, yet their quantitative estimates present several interesting puzzles.[1] These studies have consistently found the weight governments place on social welfare to be many times that which they place on lobbying effort, while numerous estimates indicate that the deadweight losses caused by trade distortions are several orders of magnitude larger than lobbying expenditures.[2] This raises the question: how could governments value social welfare so highly, yet grant so much protection at such a low price?

Attempts to more fully model the lobbying process have been successful in reducing the parameter estimates for the government's welfare considerations somewhat, but the smallest estimates still indicate that the U.S. government, for instance, values social welfare about twenty times as much as contributions.[3] Taking a different approach, Gawande and Hoekman (2006) suggest that the PFS model can be reconciled with the empirical results by acknowledging the complexity of the policy-making process and the uncertainty that arises from it.

I demonstrate how political uncertainty can be incorporated with PFS-style lobbying into a model of trade agreements in which the executive branches of the governments set their tariff levels in anticipation of political pressure upon the legislatures. The structure is similar to that of Bagwell and Staiger (2005), with two main changes: political economy weights are endogenously determined and power over the policy-making process is modeled as shared between executive and legislative branches of the government.[4] Thus, in place of Bagwell and Staiger's unitary government with differing ex-ante and ex-post preferences, this model has two branches of government with differing and—in the case of the legislature—endogenous preferences.

[1]cfr. Goldberg and Maggi 1999, Gawande and Bandyopadhyay 2000, Mitra, Thomakos and Ulubasoglu 2002, McCalman 2004.

[2]Feenstra (1992) assembles estimates from the mid-1980s that place a floor of \$8 billion per year on the deadweight losses from protection, whereas Bombardini and Trebbi (2009) find that total annual lobbying expenditure on trade issues in 1999-2001 when this data first became available was about \$200 million.

[3]cfr. Gawande, Krishna and Robbins 2006, Mitra, Thomakos and Ulubasoglu 2006, Bombardini 2008, Gawande, Krishna and Olarreaga 2005, among others.

[4]Grossman and Helpman's (1995) model of trade agreements with endogenous lobbying considers "Trade Wars" and "Trade Talks" separately, whereas this approach allows the "Trade Talks" to be influenced by the desire to avoid a "Trade War." Thus here the "Trade War" can be viewed as a crucial subgame.

This model generalizes the unitary government of the PFS setup in three ways. First, the legislature itself is assumed to be non-unitary, with the preferences of the median legislator modeled flexibly to take into account realistic lobbying strategies such as those that target committee chairs. At the very least, in legislatures with majority rules, it is clear that lobbies need win over no more than half the legislators. Since it appears that far fewer than half of legislators are actually lobbied by any given industry, this modification has the potential to provide an explanation for a significant portion of the puzzle.

Modeling power over trade policy as shared between the executive and legislative branches introduces a new dynamic and can resolve the puzzle altogether. Consider first the case of political certainty, and assume that the executive is more pro-social than the legislature, as has been the case throughout the post-war period in the United States.[5] For any trade-agreement tariff set by the executives, the lobby knows the effort level required to induce a trade disruption: it must shift the preferences of the median legislator just far enough so that he will choose the trade-war tariffs instead of the trade agreement tariffs. Because the lobby will pay this price so long as its benefit outweighs the cost, a trade war will occur unless the executive sets the trade-agreement tariff so that lobbying is not worthwhile.

Thus, any trade agreement under political certainty involves a tariff high enough to disengage the lobby completely; that is, even welfare-maximizing executives set high tariffs and induce zero contributions in equilibrium. Careful modeling of the political process demonstrates that, while tariff-setting behavior is closely linked to deadweight-loss calculations, the parameter on welfare-mindedness does not necessarily figure into those calculations in the way previously assumed. In fact, when the government is non-unitary, we must think more carefully about *whose* welfare weight is being measured. What is clear is that the earlier-cited estimates and stylized facts are perfectly consistent in this model. The "case of the missing contributions," as Gawande and Hoekman (2006)

[5]The commonly-made assumption that the executive is less protectionist than the legislature is a special case of the finding that susceptibility to special interests generally declines with the size of one's constituency. One simple illustration from the realm of trade policy is the following: a legislator whose district has a large concentration of a particular industry does not take into account the impact of tariffs on the welfare of consumers in other districts, while the executive, whose constituency encompasses the whole country, will internalize these diffuse consumption effects. For a detailed argument, see Lohmann and O'Halloran (1994).

refer to it, turns out to be an equilibrium phenomenon arising from the agenda-setting power of the executive branch.

The fact that *zero* contributions are predicted in equilibrium is not particularly satisfying. The addition of political uncertainty provides the missing realism, smooths the results and delivers additional intuition. I assume all actors are uncertain about the weight the median legislator places on the profits of the lobbying industry and that this is attributable to the complexity of the legislative process. Imagine, for instance, that the small number of lobbied legislators cannot deliver the votes of their non-lobbied colleagues with certainty. However, I show that the legislatures break trade agreements with a higher probability and set higher trade war tariffs when lobbying effort increases. Because the lobbies respond to higher trade agreement tariffs by decreasing effort (and therefore the probability the agreement will be broken), the executives must trade off the level of social welfare derived while a trade agreement is in force—since they prefer low tariffs—and the chances that the agreement will be broken.

The executives therefore set trade-agreement tariffs above their most-preferred levels in order to discourage lobbying activity and the accompanying probability of abrogation. Thus the executives use trade agreements to change the incentives for lobbying activity and the legislatures' ability to withstand political pressure. One contribution of this model, therefore, is the idea that trade agreements can act as a kind of domestic political commitment device.

Although this is not the first paper to derive trade disruptions along the equilibrium path without asymmetric information, to the best of my knowledge it is the first to model lobbying specifically aimed at derailing trade agreements. Although this kind of politically-driven failure is commonplace, prior models have not allowed exploration of the endogenous dynamics behind them. Careful consideration of the political process seems to be important for understanding why governments so often fail to cooperate.

Another outcome of this setup is that trade-agreement tariffs, lobbying effort and the probability of trade disruptions vary in nuanced ways with the amount of political uncertainty present: they are influenced strongly by lobbying incentives, but not in the straightforward ways predicted by models with unitary governments. In particular, this model can explain the approximately 15% of sectors that receive protection in spite of apparently putting forth no lobbying effort.

Ultimately, it is the addition of a rich structure of government that resolves the main empirical puzzle surrounding the PFS model: in this framework, even a welfare-maximizing executive and lobbying effort at zero can be consistent with high tariff levels.

The foundations of this work rest on Grossman and Helpman's (1994) 'Protection for Sale' and their 1995 paper 'Trade Wars and Trade Talks.' The extensive literature that explores the empirical implications of the PFS model is well-summarized by Gawande and Magee (2011). After the first generation of work consistently found very high estimates for the government's welfare-mindedness (see footnote 1 above), a second generation has explored several modifications of the model and estimating techniques. Most of these have focused on an improved accounting of the lobbying process, such as including lobbying by up- and downstream firms (Gawande, Krishna and Olarreaga 2005, Paltseva 2011), better identifying which firms in an industry lobby (Bombardini 2008, Gawande and Magee 2011), taking into account lobbying by foreign interests (Gawande, Krishna and Robbins 2006) or changing the way in which organized sectors are determined (Mitra, Thomakos and Ulubasoglu 2006). Indeed, Imai, Katayama and Krishna (2008) argue that many of the classification schemes present serious challenges for the validity of results.

Another related literature considers the impact of exogenously-determined political uncertainty on the potential for trade cooperation. These studies (e.g. Feenstra and Lewis 1991, Milner and Rosendorff 2001, Bagwell and Staiger 2005, Beshkar 2010) derive various implications for the design of international trade agreements using a Baldwin (1987)-style government welfare function with exogenous shocks to the political-pressure parameter.

Several papers have studied the impact of executive/legislative interactions on international agreements in the case of exogenously-determined preferences. Mansfield, Milner and Rosendorff (2000) construct a model in which democracies trade more with each other because the domestic legislature's influence requires the foreign government to make deeper tariff cuts to ensure its cooperation. Dai (2006) typifies a second approach, which, in line with the Schelling conjecture, argues that a country's legislative constraint assists its own executive in maintaining higher barriers. The general approach I employ here will permit me to speak to outstanding questions such as these concerning the effects of domestic institutions on trade policy.

The above work is patterned on the model of Milner and Rosendorff (1997), which explores how uncertainty can affect trade policy and the probability of ratification failure when political preferences, and therefore uncertainty, are exogenous; Iida (1996) presents a similar model. Related is Le Breton and Salanie (2003), which studies lobbying when the lobby is uncertain about the preferences of a unitary decision maker. Le Breton and Zaporozhets (2007) go a step further and replace the unitary decision maker with a legislature with multiple actors. Song (2008) presents a model in which policy-making power is shared between an executive and a unitary legislature where lobbies are able to endogenously influence the political preferences of the legislature in a context of unilateral policy making with no uncertainty.

I will begin by describing the model in detail. Section 2.2 then presents the main results. In Section 2.3, an extended example demonstrates these results as well as the impact of changing the institutional environment. Section 2.4 explores the connection between the model under consideration here and the 'Protection for Sale' theoretical framework and includes in-depth discussions of the roles of uncertainty and the separation-of-powers structure. Several extensions of the model are explored in Section 2.5 and Section 2.6 concludes.

2.1 The Model

I employ a partial-equilibrium model with two countries: home (no asterisk) and foreign (asterisk). The countries trade two goods, X and Y, where P_i denotes the home price of good $i \in \{X, Y\}$ and P_i^* denotes the foreign price of good i. In each country, the demand functions are taken to be identical for both goods, respectively $D(P_i)$ in home and $D(P_i^*)$ in foreign and are assumed strictly decreasing and twice continuously differentiable.

The supply functions for good X are $Q_X(P_X)$ and $Q_X^*(P_X^*)$ and are assumed strictly increasing and twice continuously differentiable for all prices that elicit positive supply. I also assume $Q^*_X(P_X) > Q_X(P_X)$ for any such P_X so that the home country is a net importer of good X. The production structure for good Y is taken to be symmetric, with both demand and supply such that the economy is separable in goods X and Y. It is assumed that the production of each good requires the possession of a sector-specific factor that is available in inelastic supply and is non-tradable so that the income of owners of the specific factors is tied to the price of the good in whose production their factor is used.

For simplicity, I assume each government's only trade policy instrument is a specific tariff on its import-competing good: the home country levies a tariff τ on good X while the foreign country applies a tariff τ^* to good Y. Local prices are then $P_X = P^W_X + \tau$, $P^*_X = P^W_X$, $P_Y = P^W_Y$ and $P^*_Y = P^W_Y + \tau^*$ where a W superscript indicates world prices and equilibrium prices are determined by the market clearing conditions

$$M_X(P_X) = D(P_X) - Q_X(P_X) = Q^*_X(P^*_X) - D(P^*_X) = E^*_X(P^*_X)$$

$$E_Y(P_Y) = Q_Y(P_Y) - D(P_Y) = D(P^*_Y) - Q^*_Y(P_Y) = M^*_Y(P^*_Y)$$

where M_X are home-county imports and E^*_X are foreign exports of good X and E_Y are home-county exports and M^*_Y are foreign imports of good Y.

It follows that P^W_X and P^W_Y are decreasing in τ and τ^* respectively, while P_X and P^*_Y are increasing in the respective domestic tariff. This gives rise to a standard terms-of-trade externality. As profits and producer surplus (identical in this model) in a sector are increasing in the price of its good, profits in the import-competing sector are also increasing in the domestic tariff. This economic fact, combined with the assumptions on specific factor ownership, is what motivates political activity.

I next describe the politically-relevant actors. In order to focus attention on protectionist political forces, I assume that only the import-competing industry in each country is politically-organized and able to lobby and that it is represented by a single lobbying organization. Each country's government is composed of two branches: an executive who can conclude trade agreements and a legislature that has final say on trade policy. In summary, the political process is modeled as involving three players in each country: the lobby, the executive, and the legislature.

The timing is as follows. First, the executives set trade policy cooperatively within the context of an international agreement. I assume that the agreement takes the form of tariff caps and that an external authority can ensure enforcement of the agreement in the one-shot game analyzed here.[6] After the trade agreement is concluded, the lobbies attempt to persuade the legislators in their respective countries to break the trade agreement. Next, the uncertainty about the the median legislator's identity is resolved and the legislatures decide whether to abide by the agreement or to provoke a trade war. In the event that the trade agreement does not remain in force, there is a final stage of lobbying and voting to set the trade-war tariffs. Once all political decisions are taken, producers and consumers make their decisions.

Although uncertainty is present, information about it is not asymmetric, so the appropriate solution concept is subgame perfect Nash equilibrium. As this game is solved by backward induction, it is intuitive to start by describing the incentives of the legislatures, whose decisions I model as being taken by a median legislator. As the economy is fully separable and the economic and political structures are symmetric, I focus here on the home country and the X-sector. The details are analogous for Y and foreign.

The welfare function of the home legislature is

$$W_{ML} = CS_X(\tau) + CS_Y(\tau^*) + \gamma(e,\theta) \cdot PS_X(\tau) + PS_Y(\tau^*) + TR(\tau) \tag{2.1}$$

where CS is consumer surplus, PS is producer surplus, $\gamma(e,\theta)$ is the weight placed on producer surplus (profits) in the import-competing industry, e is lobbying effort, and TR is tariff revenue.[7] Here, the weight the median legislator places on the profits of the import-competing industry, $\gamma(e,\theta)$ is affected by the level of lobbying effort[8] and the random variable θ.

I allow for the possibility of this uncertainty, which can be interpreted as uncertainty over the identity of the median legislator, in order to model the idea that the lobbying and voting "game" that goes on within legislatures is often complex enough that none of the actors know precisely what the outcome will be—that is, no one can

[6]Work in progress extends the analysis to a repeated-game context to remove this restriction.

[7]Note that, as it is conventional in the literature to neglect balanced trade for the purposes of trade revenues, I do so as well in order to maintain close comparability to previous results.

[8]The standard PFS modeling would specify $W_{ML} = C + aW$, but as will be seen when we come to the preferences of the executive, this is not sufficiently general for the purposes of this model. Although complex, an isomorphism can be made between the two forms in a special case as discussed in Section 2.4.1.

predict exactly which legislator will be the median and therefore what weight will be placed on import-industry profits at the time of the vote. Thus one way we can think about θ is as an additional influence on the median legislator's preference about which all parties—executives, legislators and lobbies alike—are uncertain.

One can conceptualize this uncertainty as being a result of the lobby's strategic choice to engage with only some key members of the legislature who wield significant influence but who cannot deliver policy decisions without the cooperation of others within the legislature, which is in turn uncertain. Different assumptions on $\gamma(\cdot,\cdot)$ should correspond to different institutional features and will affect optimal lobbying- and tariff-setting behavior.

I make the following assumptions on the weight on import-industry profits:

Assumption 1. $\gamma(e,\theta)$ *is increasing and concave in* e *for every* $\theta \in \Theta$.

Assumption 2. $\gamma(e,\theta) \geq \gamma_E \geq 1 \ \forall\theta$.

Assumption 1 formalizes the intuition that the legislature favors the import-competing industry more the higher is its lobbying effort, but that there are diminishing returns to lobbying activity. It allows the political uncertainty inherent in θ to be correlated in interesting ways with the level of e, ruling out only (i) that higher effort levels make lower γ-weights more likely and (ii) that effort and uncertainty are correlated in such a way that increasing lobbying effort changes the structure of uncertainty so that higher γ-weights become more likely at an accelerating pace.

Assumption 2 ensures that $\tau^a < \tau^n$, and more generally, that the legislature's incentives are more closely aligned with the lobby's than are those of the executive. This is not essential but simplifies the analysis and matches well the empirical findings that politicians with larger constituencies are less sensitive to special interests (See Destler 2005 and footnote 5 above).

Given its expectations and the legislature's preferences, the home lobby chooses its lobbying effort (e_b to influence the break decision and e_n to influence the trade war tariff) to maximize the welfare function:

$$\Pr\left[\text{TradeWar}(\boldsymbol{\tau^a}, e_b, e_n)\right]\left[\pi(\tau^{tw}) - e_n\right] + \Pr\left[\text{TradeAgreement}(\boldsymbol{\tau^a}, e_b, e_n)\right]\pi(\tau^a) - e_b \tag{2.2}$$

where $\pi(\cdot)$ is the current-period profit and τ^a (τ^{tw}) is the home country's tariff on the import good under a trade agreement (war). I use the convention throughout of representing a vector of tariffs for both countries (τ, τ^*) as a single bold $\boldsymbol{\tau}$.

I assume the lobby's contribution is not observable to the foreign legislature. The implication is that the lobby can directly influence only the home legislature, and so the influence of one country's lobby on the other country's legislature occurs only through the tariffs selected.[9]

In the first stage, the executives choose the trade agreement tariffs $\boldsymbol{\tau^a} = (\tau^a, \tau^{*a})$ via a negotiating process that I assume to be efficient. It therefore maximizes the expected joint payoffs:

$$W_E(\boldsymbol{\tau^a}) + W_E^*(\boldsymbol{\tau^a}) = \Pr\left[\text{TradeWar}(\boldsymbol{\tau^a})\right]\left[W_E(\boldsymbol{\tau}^{tw}) + W_E^*(\boldsymbol{\tau}^{tw})\right] + \Pr\left[\text{TradeAgreement}(\boldsymbol{\tau^a})\right]\left[W_E(\boldsymbol{\tau^a}) + W_E^*(\boldsymbol{\tau^a})\right] \quad (2.3)$$

I model the executives' choice via the Nash bargaining solution where the disagreement point is the executives' welfare resulting from the Nash equilibrium in the non-cooperative game (i.e. in the absence of a trade agreement) between the legislatures.

The home executive's welfare is specified as follows:

$$W_E = CS_X(\tau) + CS_Y(\tau^*) + \gamma_E \cdot PS_X(\tau) + PS_Y(\tau^*) + TR(\tau)$$

Note that this is identical to the welfare function for the legislature aside from the weight on the profits of the import industry, which is not a function of lobbying effort. This construction permits me to integrate the influence of the import-competing industry on negotiations prior to the formation of the trade agreement while also reflecting the idea that the mechanisms through which political considerations influence the executives' preferences appear to be quite different from those at the legislative level (the electoral college and primary/caucus systems for instance).

This assumption does *not* require that the executives are not lobbied; only that their preferences are not directly altered in a significant way by lobbying over trade—that they do not sell protection in order to finance their re-election campaigns. In the case

[9] cfr. Grossman and Helpman 1995, page 685.

of the post-war United States, where the Congress has consistently been significantly more protectionist than the President, this seems to reasonably reflect the political reality. For trade policy, where there are concentrated benefits but harm is diffuse, there are good reasons for this to be the case. Because the President has the largest constituency possible, delegating authority to the executive branch may simply be a mechanism for "concentrating" the benefits since consumers seem unable to overcome the free-riding problem. In fact, a strong argument can be made that power over trade policy has been delegated to the executive branch precisely *because* it is less susceptible to the influence of special interests (Destler 2005).

Therefore, in line with both the theoretical and empirical literature, I will assume that $\gamma_E \leq \gamma(e,\theta)$ for all realizations of θ.[10] That is, even for the least favorable outcome of the lobbying process, the legislature will be at least slightly more protectionist than the executive. This does not fully explain why the executive branch does not seem to be influenced toward protectionism in the same way that the legislature is, but one plausible rationale is that the President faces a more complex electoral calculus over a wider-range of policy areas.

Although the political process here matches most closely that of the United States in the post-war era, I believe the model or one of its extensions is applicable for a broad range of countries for which authority over the formation and maintenance of trade policy is diffuse and subject to political pressure either at home or in a trading partner.[11]

2.2 Main Results

To understand how the executives optimally structure trade agreements, we must first examine the incentives of the lobbies and how the legislatures make decisions regarding breach of the trade agreement, including how trade-war tariffs are set. The symmetric structure of the model permits restriction of attention to the home country.

[10]Note that this includes in particular the special case of a welfare-maximizing executive ($\gamma_E = 1$).

[11]In particular, the binary decision by the legislature about whether to abide by or break the trade agreement is modeled on the "Fast Track Authority" that the U.S. Congress granted to the Executive branch almost continuously from 1974-1994 and then again as "Trade Promotion Authority" from 2002-2007.

2.2.1 Trade-War Tariffs

In the event that the trade agreement is broken, the legislature sets its tariff τ unilaterally by maximizing Equation 2.1 given τ^*. Because there are no interactions between the home and foreign tariffs, the home and foreign trade-war tariffs are independent and the home country's tariff in a trade war maximizes weighted home-country welfare in the X-sector only. The foreign legislature's decision problem is analogous, and unilateral optimization leads to what I refer to as the Nash tariffs as the solution to the following first order condition:

$$\frac{\partial CS_X(\tau)}{\partial \tau^n} + \gamma(e_n, \theta_n) \cdot \frac{\partial PS_X(\tau)}{\partial \tau^n} + \frac{\partial TR(\tau)}{\partial \tau^n} = 0$$

Conditions under which the second order condition is satisfied are contained in Lemma 6 in the Appendix.

In accordance with intuition, because profits are increasing in the tariff, trade war tariffs are increasing in the weight attached to the profits of the import-competing industry, as the following result shows:

Lemma 1. *If prices are linear in tariffs, $\tau^n(\gamma(e_n, \theta_n))$ is increasing in $\gamma(e_n, \theta_n)$.*

Proof: See the Appendix.

In the event of a trade war, the lobby chooses its effort e_n given this tariff-setting behavior by maximizing its profits net of effort: $\pi\left(\tau^n\left(\gamma\left(e_n, \theta_n\right)\right)\right) - e_n$. (Note this is Equation 2.2 simplified by the resolution of uncertainty over the legislature's decision on the trade agreement). This implies a first order condition of

$$\frac{\partial \pi(\tau^n(\gamma(e_n, \theta_n)))}{\partial e_n} = 1 \qquad (2.4)$$

That is, at this stage, the lobby chooses the level of effort that equates its expected marginal increase in profits with its marginal payment. Further details about the political weighting function γ and assumed structure of uncertainty are provided in the next section.

2.2.2 To Break or Not to Break?

With the trade-war behavior of both the lobby and the legislature specified, we can proceed to analyze their interaction regarding the legislature's decision to uphold or break the trade agreement. I assume that the form of the political weighting function is the same in both cases, but the model could be extended to allow the two to differ. What is in principle different are the lobbying efforts toward influencing the break decision, denoted e_b at this stage to distinguish them from those that are made to influence the Nash tariff, e_n. The realization of uncertainty about the median legislator's identity will similarly be denoted θ_b so as to distinguish it from θ_n at the later stage.

I also assume that each legislature will have the opportunity to break the agreement with probability less than one-half, and that the probability that both legislatures have the opportunity is vanishingly small. This allows me to focus on the interaction between the domestic actors at this stage; I discuss an extension to the case in which both legislatures could potentially vote to break the agreement at the same time in Section 2.5.1.

The legislature will break the agreement and set the tariff at τ^n if the median legislator's utility from the Nash tariffs[12] is higher than his utility from the trade agreement tariffs, i.e. if

$$W_{ML}(\boldsymbol{\tau^n}, \gamma(e_b, \theta_b)) > W_{ML}(\boldsymbol{\tau^a}, \gamma(e_b, \theta_b)) \tag{2.5}$$

Recall that a bold $\boldsymbol{\tau}$ represents a vector of tariffs for both countries (τ, τ^*).

The legislature's decisions are stochastic, so the outcome of the vote on whether or not to break the trade agreement, as well as the level at which the tariff will be set in the case of a disruption, is not known to *any* player until the uncertainty over the identity of the median legislator is resolved at each stage—that is, until the moment the vote takes place.[13] I represent the probability that the home legislature breaks the trade

[12]Note again the assumption of external enforcement: the countries can choose to abide by the terms of the agreement and take advantage of the external enforcement or can exit the agreement, but they cannot "cheat."

[13]Although we might think of the trade-war tariff as being set in the legislation before a vote takes place, the results of this slight simplification accord closely with intuition.

agreement and sets the tariff at τ^n as:

$$b(e_b, \boldsymbol{\tau^a}, \boldsymbol{\tau^n}) = \mathbb{E}_{\gamma|e_b} \mathbf{1}[W_{ML}(\boldsymbol{\tau^n}, \gamma(e_b, \theta_b)) > W_{ML}(\boldsymbol{\tau^a}, \gamma(e_b, \theta_b))]$$
$$= \Pr[W_{ML}(\boldsymbol{\tau^n}, \gamma(e_b, \theta_b)) > W_{ML}(\boldsymbol{\tau^a}, \gamma(e_b, \theta_b)) | e_b] \quad (2.6)$$

We are now in a position to examine the legislature's decision more closely. Of central concern is how the probability that the legislature will break the trade agreement varies with lobbying effort:

Result 1. *The probability that the legislature breaks a trade agreement is increasing and concave in* e_b.

Proof: See the Appendix.

Lobbying affects only the weight the legislature places on the profits of the import-competing industry. These profits are higher in a trade war than under a trade agreement, so given Assumption 1, the expectation of γ is increasing in lobbying effort, implying that the legislature becomes more favorably inclined toward the high trade-war tariff and associated profits as lobbying increases and therefore more likely to break the trade agreement.

The probability of a trade war, just like the trade-war tariffs, is increasing in γ, so it's the natural assumption that γ is increasing in lobbying effort that implies that the lobby can increase both its trade war tariff and the probability of a trade war by raising its lobbying effort.

Turning to the effects of the trade-agreement tariffs on the probability that the agreement will be abrogated, it is straightforward that the legislature always prefers lower levels of the foreign tariff:

Lemma 2. *The probability the legislature breaks a trade agreement is increasing in* τ^{*a}.

Proof: See the Appendix.

Because the home country is an exporter of good Y, the lower world price for Y has a negative effect on producer surplus that is larger than the positive effect on consumer surplus. Thus the net effect of an increase in foreign trade-agreement tariffs on legislative (and home country) welfare is negative, leading to an increased probability that the trade agreement will be broken.

The relationship between the home country's trade-agreement tariff τ^a and the legislature's probability of breaking the agreement is slightly more complex because τ^a interacts directly with $\gamma(e_b, \theta_b)$. The intuition is as follows. For any given level of e_b, we can think of the expected median legislator as having a preferred tariff that can be found in the same way as τ^n. Call this tariff $\mathbb{E}\left[\tau^{e_b}\right]$, and notice that, in expectation, the median legislator's welfare under the trade agreement increases in τ to the left of $\mathbb{E}\left[\tau^{e_b}\right]$ and decreases in τ to the right of it. The break probability therefore decreases up to $\mathbb{E}\left[\tau^{e_b}\right]$ because the trade agreement tariff is coming ever closer to the expected median legislator's preferred level. However, it continues to decrease thereafter because, as the tariff rises further, progressively more protectionist median legislators will also prefer the trade agreement—that is, fewer and fewer realizations of $\gamma(e_b, \theta_b)$ will lead to abrogation. Even very high tariffs, although not pleasing to the median legislators that result from low realizations of $\gamma(e_b, \theta_b)$, are preferable to a trade war, so the probability of breaking either becomes zero or continues to decrease.

Lemma 3. *Holding lobbying effort constant, the probability the legislature breaks a trade agreement is weakly decreasing in τ^a.*

Proof: See the Appendix.

The intuition for the above results is quite straightforward. When the lobby increases effort, the median legislator will have a higher weight on the import-competing industry's profits and prefer a higher tariff, making it more likely to break the trade agreement. On the other hand, a higher tariff for the import-competing industry (at least for those below the expected median legislator's most-preferred level) increases the legislature's payoffs under the trade agreements and makes it less likely to break that agreement, while a higher tariff on home country exports in the trading partner reduces the legislature's payoffs from the agreement and makes it more likely to break the agreement.

Lobby

The lobby chooses its level of effort e_b as a function of $\boldsymbol{\tau^a}$, given the implications of that choice on the legislature's probability of breaking the agreement, $b(e_b, \boldsymbol{\tau^a}, \boldsymbol{\tau^n})$. I will henceforth suppress the Nash tariffs in the expression of the break probability to simplify notation.[14] Recalling that if the lobby has the chance to act, there is no chance for the foreign legislature to break the agreement, the lobby's decision is

$$\max_{e_b} b(e_b, \boldsymbol{\tau^a}) \left[\pi(\tau^n) - e_n\right] + \left[1 - b(e_b, \boldsymbol{\tau^a})\right]\pi(\tau^a) - e_b$$

The lobby simply maximizes probability-weighted profits net of effort. From its first order condition, we can see that the lobby balances the cost of an extra dollar of lobbying expenditure with the higher profits from a trade war weighted by the increase in the probability of receiving the higher profit level that expenditure induces:

$$\frac{\partial b(e_b, \boldsymbol{\tau^a})}{\partial e_b} \left[\pi(\tau^n) - e_n - \pi(\tau^a)\right] = 1 \tag{2.7}$$

The second order condition is satisfied given Assumption 2 and Result 1. However, the solution is not guaranteed to be interior, for which we need

$$\frac{\partial b(0, \boldsymbol{\tau^a})}{\partial e_b} \left[\pi(\tau^n) - e_n - \pi(\tau^a)\right] > 1. \tag{2.8}$$

For instance, if the executives were to set $\tau^a = \tau^n$, there would be no incentive for the lobby to make a positive contribution. We will see that there are some cases in which it may be in the executives' joint interest to set trade agreement tariffs so as to completely disengage the lobby. The following results thus only hold when Condition 2.8 is satisfied, that is, when the marginal impact of the first lobbying dollar on the break probability is sufficiently high to make engaging in the political process worthwhile for the lobby.[15]

We can now proceed to the central result concerning the lobby's behavior.

Result 2. *When the trade agreement remains in force with positive probability, lobbying effort is decreasing in the home-country trade agreement tariff.*

[14]Since the decisions over the Nash tariffs are taken in the final period, they are constants from the point of view of earlier stages. Similarly, because the lobbying effort made in that period does not interact with other decision variables in earlier periods, I will suppress the dependence of the Nash tariffs on e_n.

[15]This is primarily a condition on the curvature of the political weighting function but also involves the production and demand structure.

Proof: See the Appendix.[16]

Raising the trade agreement tariff decreases the benefit of inducing a break in the trade agreement by raising profits under the trade agreement. This is in many ways the key result of this paper: the executives can reduce lobbying effort by setting higher tariffs in their trade agreement. We will see in the next section how this shapes the executives' joint decision.

Note that, because the foreign tariff does not impact profits in the industry that is represented by the lobby, changing τ^{*a} does not change the optimal choice of lobbying expenditure.

Corollary 1. *Lobbying effort does not respond to the foreign-country trade agreement tariff.*

Proof: See the Appendix.

2.2.3 The Trade Agreement

The executives choose trade agreement tariffs $\boldsymbol{\tau^a} = (\tau^a, \tau^{*a})$ to maximize Equation 2.3. I call the maximized joint value $MV(\boldsymbol{\tau^a})$ and write the division of surplus according to the Nash bargaining solution as

$$V_E(\boldsymbol{\tau^a}) + t = W_E(\boldsymbol{\tau^n}) + \frac{1}{2}\left(MV(\boldsymbol{\tau^a}) - W_E(\boldsymbol{\tau^n}) - W_E^*(\boldsymbol{\tau^n})\right)$$

$$V_E^*(\boldsymbol{\tau^a}) - t = W_E^*(\boldsymbol{\tau^n}) + \frac{1}{2}\left(MV(\boldsymbol{\tau^a}) - W_E(\boldsymbol{\tau^n}) - W_E^*(\boldsymbol{\tau^n})\right)$$

where expectation operators have been dropped for convenience and t is an intergovernmental transfer.[17] The solution to this system of equations determines the tariffs $\boldsymbol{\tau^a}$ that will be set in the trade agreement.

I simplify the problem by assuming that political constraints prevent the executives from choosing asymmetric tariffs. This restricts attention to symmetric solutions in

[16] See Section 2.4.3 for a detailed account of behavior when the lobby is able to induce a trade war with probability 1.

[17] While direct monetary transfers have to date rarely been used in practice, it seems appropriate to interpret linked concessions on non-trade issues as indirect transfers (Klimenko, Ramey and Watson 2008, Maggi and Staiger 2011). Here, transfers do not occur as long as the full set of symmetry assumptions are maintained but are an important consideration in the environment of Section 2.5.1.

the fully symmetric environment under consideration, which means that the trade agreement tariffs for which we're looking are simply those that maximize the joint welfare of the executives.[18] In order to solve for them, I represent the probability that the trade agreement will be broken as $B(\boldsymbol{\tau^a}) = b(e(\boldsymbol{\tau^a}), \boldsymbol{\tau^a})$ where $e(\boldsymbol{\tau^a})$ is the best response function implicit in Equation 2.7, the lobby's first order condition.

The problem then becomes to maximize the following modified version of Equation 2.3:

$$\{o \cdot B(\boldsymbol{\tau^a}) + o^* \cdot B^*(\boldsymbol{\tau^a})\} \boldsymbol{W_E}(\boldsymbol{\tau^n}) + \{1 - o \cdot B(\boldsymbol{\tau^a}) - o^* \cdot B^*(\boldsymbol{\tau^a})\} \boldsymbol{W_E}(\boldsymbol{\tau^a}) \quad (2.9)$$

where I have written the sum of the home and foreign executives' utilities as $\boldsymbol{W_E}(\cdot)$. Here o and o^* are the probabilities that the home and foreign legislatures, respectively, will have the opportunity to break the trade agreement; these are assumed to be mutually exclusive events that are realized after the conclusion of the trade agreement. The case in which both legislatures can potentially break the agreement is treated in Section 2.5.2. One can think of o and o^* as being determined by events beyond the executives' or legislatures' control that determine whether or not the legislatures will be willing to consider the lobby's request; for example, they would be affected by the occurrence of an economic crisis that diverts the legislature's attention from less-pressing matters.

Of central concern is how the break probabilities behave as a function of the trade agreement tariffs given the lobby's optimal response. Begin by examining the impact of $\boldsymbol{\tau^a}$ on $B(\boldsymbol{\tau^a})$. Two effects are present: the indirect effect through the impact on the lobby's choice of e_b, which is negative by Results 1 and 2 and Corollary 1, and the direct effect on the legislature's break decision. Lemma 3 tells us that the latter is also negative. However, because we've assumed that τ^a must equal τ^{*a}, any increase in τ^a will be accompanied by an equal increase in τ^{*a}. This does not affect the lobby's effort, but Lemma 2 shows that τ^{*a} has a positive direct impact on the legislature's probability

[18] I have assumed that $W_E(\tau) = W_E^*(\tau)\ \forall \tau$; further in a symmetric equilibrium, $\tau^n = \tau^{*n}$ and $\tau^a = \tau^{*a}$, so

$$\begin{aligned} V_E(\boldsymbol{\tau^a}) + t &= W_E(\boldsymbol{\tau^n}) + \frac{1}{2}\left(MV(\boldsymbol{\tau^a}) - W_E(\boldsymbol{\tau^n}) - W_E^*(\boldsymbol{\tau^n})\right) \\ &= W_E(\boldsymbol{\tau^n}) + \frac{1}{2}\left(MV(\boldsymbol{\tau^a}) - 2W_E(\boldsymbol{\tau^n})\right) = \frac{1}{2}\mathbb{E}MV(\boldsymbol{\tau^a}) \end{aligned}$$

of breaking the trade agreement. The following result addresses the combined impact of τ^a and τ^{*a} on $b(e, \boldsymbol{\tau^a})$ when the two are constrained to be equal:

Lemma 4. *Holding lobbying effort constant, the probability the legislature breaks a trade agreement is weakly decreasing in $\boldsymbol{\tau^a}$ (i.e. $\frac{\partial b(e,\boldsymbol{\tau^a})}{\partial \tau^a} + \frac{\partial b(e,\boldsymbol{\tau^a})}{\partial \tau^{*a}} \leq 0$).*

Proof: See the Appendix.

When any change in one of the trade agreement tariffs must be met with an equal change in the other (as in a symmetric agreement), the negative impact on the legislature's propensity to break the trade agreement of raising the home tariff always outweighs the positive impact of raising the foreign tariff. This fully describes the direct effect of the trade agreement tariffs on the total break probability. Recalling the definition of the total break probability, $B(\boldsymbol{\tau^a}) = b(e_b(\boldsymbol{\tau^a}), \boldsymbol{\tau^a})$, we must combine this direct effect with the indirect effect through the lobbying decision. The fact that the indirect effect of raising tariffs is a reduction in lobbying (see Result 2) and therefore a reduction in the propensity of the legislature to break the agreement (Result 1) leads to the following key result: the overall probability that the trade agreement will be broken is decreasing in the trade agreement tariffs.

Result 3. *The total probability that the trade agreement will be broken is decreasing in $\boldsymbol{\tau^a}$ (i.e. $\frac{\partial B(\boldsymbol{\tau^a})}{\partial \tau^a} + \frac{\partial B(\boldsymbol{\tau^a})}{\partial \tau^{*a}} \leq 0$).*

Proof: See the Appendix.

That is, when the executives raise the trade agreement tariffs the legislature becomes less likely to abrogate the agreement, for three reasons. First, the legislature prefers a higher domestic tariff; second, the higher tariff discourages lobbying; finally, the lower lobbying effort reduces the legislature's preferred tariff further. Thus, beyond promising a lower tariff from the trading partner, we can think of the trade agreement as a sort of political commitment device that can be used to relieve political pressure and allow the legislature to maintain a lower tariff than it otherwise would.

We are now prepared to examine the executives' optimal choice of trade agreement tariffs. Because joint executive welfare is decreasing in trade agreement tariffs for $\boldsymbol{\tau^a}$ above the executives' preferred tariffs, we have the following fundamental feature of the executives' problem:

Lemma 5. *The executives face the following trade off when choosing τ^a: higher tariffs decrease the probability that the trade agreement will be broken, but they also decrease welfare when the agreement is in force.*

I define the tariff the executives prefer in the absence of legislative constraints as τ^E and show in the Appendix that it is never optimal in a symmetric equilibrium for the executives to choose $\tau^a < \tau^E$ if there is a positive probability that the legislatures will have an opportunity to break the trade agreement. This is the result of an envelope-like result: there are only second-order losses from raising tariffs slightly from the most-preferred level yet there are first-order gains in reducing the probability that the agreement will be broken.

The executives will also always choose $\tau^a < \tau^n$ unless the legislature breaks even agreements with tariffs very close to the Nash level with certainty, in which case the problem is not interesting so I will ignore this case. Thus there is an interior solution in all cases of interest, but this solution may be of two different forms. First, it could be at the point that maximizes the concave portion of the executives' welfare function (Equation 2.9). Here, the optimal tariffs are chosen to balance concerns for trade-agreement welfare against those for discouraging lobbying activity aimed at breaking the agreement.

But there is an alternative. For some specifications, high enough tariffs can disengage the lobby sufficiently so that the probability that the trade agreement will be broken is reduced to zero. Interestingly, if this occurs at a low enough tariff level, executive welfare can be maximized at this point.

Result 4. *The executives maximize their welfare by either (a) raising tariffs sufficiently high to ensure that the trade agreement will remain in force or (b) trading off reductions in the probability that the agreement will be broken with reductions in welfare under the agreement.*

This seems to accord well with observations of trade policy politics, in that some lobbies appear to exert significant effort toward disrupting trade agreements (often in the form of some kind of individual administrative remedy such as anti-dumping duties) whereas others apparently do not contest the levels of protection granted them. The current model points to differences across industries in production and demand structure, as well as political weighting (γ), to help explain these variations.

2.3 An Example

It is instructive to examine a simple parameterization of the model economy. The fundamentals here are chosen to match those of Bagwell and Staiger (2005). Home country demand, supply and profits are given by $D(P_i) = 1 - P_i$, $Q_X(P_X) = \frac{P_X}{2}$, $Q_Y(P_Y) = P_Y$, $\Pi_X(P_X) = \frac{(P_X)^2}{4}$, and $\Pi_Y(P_Y) = \frac{(P_Y)^2}{2}$ where P_i is the price of good i in the home country market. Foreign is taken to be symmetric.

This implies Home-country imports of X and exports of Y of $M_X(P_X) = 1 - \frac{3}{2}P_X$ and $E_Y(P_Y) = 2P_Y - 1$, with foreign imports of Y and exports of X given by $M_Y^*(P_Y^*) = 1 - \frac{3}{2}P_Y^*$ and $E_X(P_X^*) = 2P_X^* - 1$. With the only trade policy instruments being tariffs on import competing goods, world prices are $P_X = P_X^W + \tau$, $P_X^* = P_X^W$, $P_Y^* = P_Y^W + \tau^*$, and $P_Y = P_Y^W$. Market clearing implies that world and home prices of X are $P_X^W = \frac{4-3\tau}{7}$ and $P_X = \frac{4+4\tau}{7}$, symmetric for Y.

2.3.1 Trade War Tariffs

The median legislator's welfare can be written as the sum of

$$W_{ML}^X(\tau, \gamma(e,\theta)) = \frac{9}{98} - \frac{5}{49}\tau - \frac{34}{49}\tau^2 + \frac{1}{98}\gamma(e,\theta)\left[8 + 16\tau + 8\tau^2\right]$$

$$W_{ML}^Y(\tau^*) = \frac{25}{98} - \frac{3}{49}\tau^* + \frac{9}{49}(\tau^*)^2$$

where $W_{ML}^X(\tau, \gamma(e,\theta))$ is the utility derived from consumer surplus, producer surplus and tariff revenues in the import-competing industry and $W_{ML}^Y(\tau^*)$ is the utility derived from consumer and producer surplus in the exporting industry.

When setting the trade-war tariff, the legislature simply maximizes $W_{ML}(\tau, \tau^*) = W_{ML}^X(\tau) + W_{ML}^Y(\tau^*)$ by choice of τ given τ^*. As there are no interactions between τ and τ^*, the legislature simply maximizes $W_{ML}^X(\tau)$ and sets the trade war tariff

$$\tau^n = \frac{8\gamma(e,\theta) - 5}{68 - 8\gamma(e,\theta)}$$

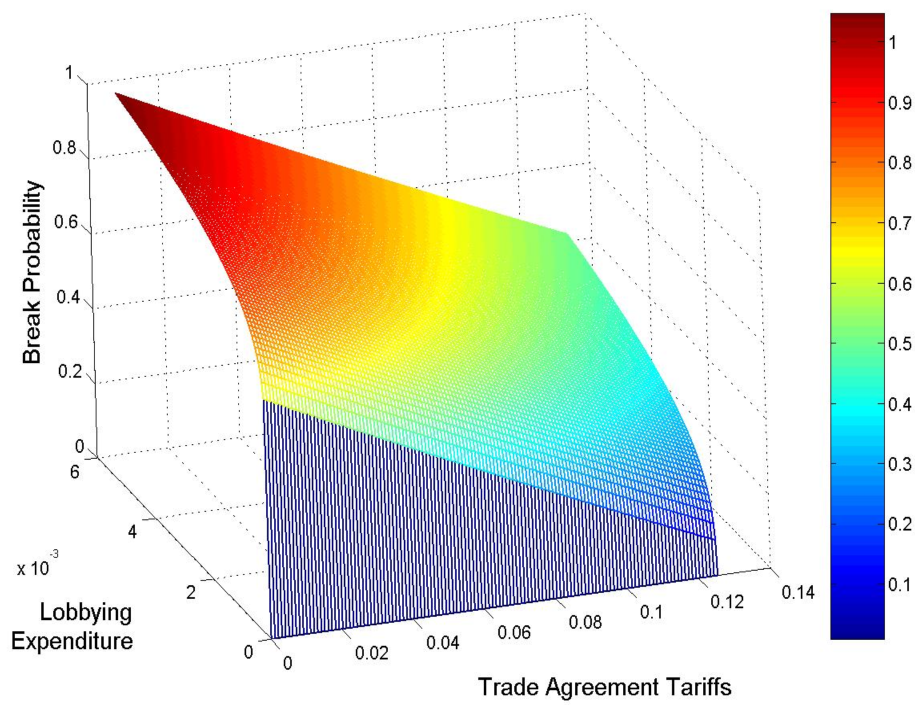

Figure 2.1: Probability trade agreement will be broken

I refer to this as the Nash tariff because it is the result of Nash equilibrium in the non-cooperative game between the legislatures. τ^n is increasing in e and the second order condition is satisfied for all realizations of $\gamma < 17/2$. Because $\gamma = 7/4$ is enough to achieve the prohibitive tariff of $1/6$ it seems reasonable to assume that this condition is satisfied.

In the event of a trade war and facing this tariff-setting behavior by the legislature, the lobby maximizes $\pi\left(\tau^n\left(\gamma\left(e_n, \theta_n\right)\right)\right) - e_n$. In order to predict the Nash tariff, the political weighting function and form of political uncertainty must be specified. Take for instance $\gamma(e, \theta) = 1.25 + C^{.2}$ with θ distributed uniformly on $[-0.25, 0.25]$. Facing this specification of the political process, the lobby maximizes its objective function at $e_n = 0.00166$, which produces a Nash tariff of 0.129.

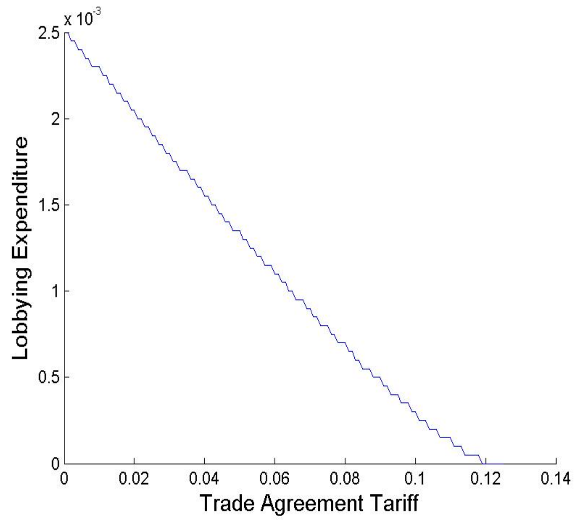

Figure 2.2: Lobbying effort

2.3.2 Break Decision

Next we move to the legislature's decision on whether or not to break the trade agreement. Figure 2.1 depicts the probability that the legislature will vote to break the trade agreement as a function of the tariff levels set in the trade agreement (with the restriction that $\tau^a = \tau^{*a}$) and the lobby's effort. It is strongly increasing in lobbying effort and decreasing in the level of tariffs set in the trade agreements.

Given the impact of lobbying on the legislature's break decision, the lobby's optimal contribution level turns out to be strongly decreasing in the trade agreement tariffs, as shown in Figure 2.2.

2.3.3 Trade Agreement

Because the probability of legislative break never reaches zero for this parameterization, the executives' welfare function is concave everywhere. Assuming that each legislature has the opportunity to vote to break the agreement with probability $\frac{1}{2}$ and that the executives are social-welfare maximizers (i.e. $\gamma_E = \gamma_E^* = 1$), they will set trade agreements tariffs of $\tau^a = \tau^{*a} = 0.078$, with lobbying expenditures of 0.0007 and a total break probability of 0.505. The expected tariff is then 0.103.

We can compare this against several different benchmarks. The most stark is the trade-war outcome itself, which can also be interpreted as the outcome that would prevail in the absence of executive involvement in trade policy *and* the absence of any effort to make a trade agreement. The tariff in the executive-formed trade agreement is about 60% of the Nash tariff of 0.129, and the expected level given the probability that the agreement will be broken is 80% of the Nash level. Lobbying expenditures in the agreement-maintenance phase are about 40% of the those in the Nash game, while expected expenditures in the executive-led trade-agreement scenario are 90% of what they would be if the legislatures set tariffs unilaterally.

Although the welfare-maximizing governments here are not able to set and maintain tariffs at zero as they would like, they are able to achieve significant reductions in tariff levels through the use of the trade agreement.

Another interesting benchmark is a scenario in which no executives are involved in trade policy, but in which the politically-susceptible legislatures can make use of a trade agreement. As in Bagwell and Staiger (2005), I find when maximizing their joint welfare in a legislature-led trade agreement ($W_{ML}(\tau, \tau^*) = W_{ML}^X(\tau) + W_{ML}^Y(\tau^*)$), the cooperative tariff levels will be set at

$$\tau^L = \frac{4\gamma(e,\theta) - 4}{25 - 4\gamma(e,\theta)}, \ \tau^{*L} = \frac{4\gamma(e^*,\theta) - 4}{25 - 4\gamma(e^*,\theta)}$$

The legislatures are able to use the agreement to internalize the terms of trade externality, making political influence more expensive for the lobbies. Indeed, I find that lobbying expenditures rise to 0.0027—60% higher than in the Nash case—while the agreement tariffs are set at 0.118, only slightly lower than the Nash tariffs of 0.129 and still higher than even the expected tariffs from the trade agreement scenario in which the executives choose the tariffs.

If one interprets the addition of a second policy-making actor as consistent with a move from autocracy to democracy, this result helps to explain the extensive empirical evidence that democracies trade more with each other than autocracies.[19]

[19]cfr. Laver and Shepsle 1999; Morrow, Siverson and Tabares 1999; Milner and Kubota 2005; Henisz and Mansfield 2006; Kono 2006; Ward and Hoff 2007; Aidt and Gassebner 2010.

2.4 Discussion

2.4.1 Relation to Grossman and Helpman's 'Protection for Sale'

The legislative welfare function employed in this paper allows significant flexibility to model a wide range of political processes within a non-unitary legislature. This is a departure from the 'Protection for Sale' welfare function, which assumes that a unitary government maximizes the sum of contributions and some fraction, a, of social welfare. The $\gamma(e,\theta)$-weighting procedure does not rule out that any given politician behaves according to the PFS specification, but it does allow for legislators to have different a-weights and for the aggregation of their preferences to follow more realistic patterns.

To see the relationship between the two forms, we can write the PFS welfare function (replacing their C with e) as

$$e + aW = e + a\left[CS_X(\tau) + CS_Y(\tau^*) + PS_X(\tau) + PS_Y(\tau^*) + TR(\tau)\right]$$

Here, just as in Equation 2.1, the relationship between e and the weight the legislature places on profits is endogenous. However, in Equation 2.1, the weight placed on profits is determined by the reduced form $\gamma(e,\theta)$, whereas in the PFS model, we must first solve for equilibrium behavior to retrieve the relationship between e and the weight the legislature places on profits.

Thus the isomorphism to the model presented in this paper can only be made once equilibrium lobbying behavior has been determined—that is, after specifying how the game will be played and so how lobbying effort will be related to the weight on profits in equilibrium. Although $\gamma(e,\theta)$ is a reduced-form of this relationship, in a model with only one lobby who makes a take-it-or-leave-it offer, the lobby's behavior in this regard is very easy to describe and so nothing is lost.

Take the case of no uncertainty and assume that the legislature chooses $\tau = 0$ in the absence of lobbying and that $\gamma(0) = 1$. In the PFS model, for any τ that the lobby desires, it must pay according to the indifference condition

$$e(\tau) = a\left[W(0) - W(\tau)\right]$$

That is, the lobby must pay for the welfare loss caused by the tariff it requests, weighted by a. In the model of this paper, the indifference condition is

$$CS_X(\tau) + CS_Y(\tau^*) + \gamma(e) \cdot \pi_X(\tau) + \pi_Y(\tau^*) + TR(\tau) = \\ CS_X(0) + CS_Y(0) + \gamma(0) \cdot \pi_X(0) + \pi_Y(0) + TR(0)$$

or

$$\gamma(e) = \frac{CS_X(0) + CS_Y(0) + \pi_X(0) + \pi_Y(0) + TR(0)}{\pi_X(\tau)} - \frac{CS_X(\tau) - CS_Y(\tau^*) - \pi_Y(\tau^*) - TR(\tau)}{\pi_X(\tau)}$$

In order to match the form of the simple PFS framework above, we must have

$$\gamma(e) = 1 + \frac{e}{a\pi_X(\tau)} \tag{2.10}$$

Choosing the political weighting function identified in Equation 2.10 aligns the political objective function used in the current model with the PFS framework. Although using the γ-weighted government welfare function is a reduced-form approach in one sense, the above calculations demonstrate that the PFS welfare function embodies a very specific restriction on the political weights: a lobby must pay exactly a-weighted welfare cost of the policy it receives.

This helps to address the questions that have been raised by empirical investigations of the PFS model about the relationship of the magnitude of welfare losses relative to the weight placed on social welfare by governments. Most studies have predicted the pattern of protection quite well but produce parameter estimates for a that seem to be up to several orders of magnitude too large for the deadweight losses produced by the protection provided.

While the empirical work is well-placed to test the cross-industry implications of the PFS model, it is unable to test the political assumption that the government is a unitary actor. In fact, it is quite clear that trade policy is shaped in a very complicated process (see, for instance, Destler 2005) involving multiple actors. This paper has presented one attempt to more fully model the political process—both by adding a second branch of government and by acknowledging the complexity of the decision-making process within the legislature.

Even without the involvement of the executive branch, the most natural extension of the PFS assumption to a non-unitary legislature would have the industry lobbying all legislators in exactly the same way, similar in spirit to Le Breton and Zaporozhets (2007). Actual behavior by lobbies tends to follow a very different pattern in which a subset of legislators receive contributions and lobbying activity in varying amounts and then participate in a complicated process involving committees and logrolling (Ansolabehere, Figueiredo and Snyder 2003). So while the amounts of protection granted may conform quite well to the cross-industry PFS predictions, the issue of the magnitudes of contributions and lobbying expenditures being too low compared to the deadweight losses inflicted by the protection provided may be explained in part by the mismatch between the unitary government assumption and the reality of the political process. One of the simplest hypotheses that emerges from the non-unitary government framework is that lobbies may only have to pay a small number of legislators for their districts' a-weighted welfare loss.

The current model demonstrates that the solution to the empirical puzzle is most likely more complex. Examining a more realistic political process and the concomitant political uncertainty reveals a nuanced relationship between the preferences of governmental actors, lobbying and tariffs. The executives' incentive to set higher tariffs than they would otherwise prefer in order to discourage lobbying activity and trade disruptions introduces a wedge between lobbying expenditures and the associated political weights on one hand and tariff levels on the other. The next section explores the effects of the separation of powers between the executive and legislative branches of the government on lobbying expenditures and the associated provision of tariffs in more detail, while Section 2.4.3 addresses the additional impact of political uncertainty.

2.4.2 Separation of Powers

Consider the model of Section 2.2 for the case of $\theta_b = 0$—that is, no uncertainty at the break stage (as will be discussed in the next section, in the case of mean-zero uncertainty, strategic behavior at the trade-war stage is not altered by uncertainty). This allows the isolation of the results that derive from the assumption that power over trade policy is shared between the executive and legislative branches from those which derive from uncertainty over the identity of the median legislator. Again, let us focus on the home lobby.

Now $\gamma(e_b)$ is deterministic, so the lobby knows the precise contribution it must make for any given trade agreement tariffs $\boldsymbol{\tau^a}$ to induce the legislature to break the agreement (see Inequality 2.5). Here, the lobby's contribution *increases* in $\boldsymbol{\tau^a}$ since the higher are the trade agreement tariffs, the larger is the political weight that is required to induce the legislature to find them unsatisfactorily low. As long as trade war profits net of the required contribution are greater than trade agreement profits, the lobby will induce a trade war; otherwise, it is not in the lobby's interest to make any contribution. Facing this behavior, the executives prefer to set the lowest trade agreement tariffs that make it prohibitively expensive for the lobby to have the agreement broken.

I return to the example of Section 2.3 to illustrate the effects of changing the environment to one without political uncertainty. Recall that, under the specification $\gamma(e, \theta) = 1.25 + e_b^2$ with θ_b distributed uniformly on $[-0.25, 0.25]$, the Nash tariff is 0.129 and the executive-led trade agreement tariff is 0.078, while if the legislatures form a trade agreement directly, they would set tariffs of 0.118 while the lobbies would spend 0.0027.

In this example, under an executive-led trade agreement with no uncertainty, the tariff level will be set at 0.106, the lobby will exert zero effort, and the agreement will remain in force with probability 1. If instead the executives were to set trade-agreement tariffs of 0.105, the lobby would contribute 0.0025 and the legislature would break the agreement with probability 1. This demonstrates both the agenda-setting power of the executives and the stark discontinuities induced when political uncertainty is not an issue.

Moreover, this case highlights the manipulation of the tariff level to discourage lobbying and therefore the disconnect that can arise between the tariffs that are chosen and the preferences of the governmental actors. In the example, the executive is a social welfare maximizer and so prefers zero tariffs, while the legislators with $\gamma(e, \theta) = 1.25 + e_b^2$ with $e_b = 0$ prefer 0.05. However, the trade agreement tariffs are set at 0.106 precisely to ensure that $e_b = 0$ so that we have a relatively free-trading legislature that upholds the trade agreement.

2.4.3 The Role of Political Uncertainty

In this model, if there were no political uncertainty, there would be no lobbying in equilibrium. Thus, if nothing else, adding the realism of political uncertainty allows for positive lobbying on the equilibrium path. If we examine more closely the assumption that there is significant uncertainty surrounding the legislative lobbying process, additional insights come to light. It is most clear in the case of mean-zero uncertainty, so I will assume this throughout this section unless otherwise noted.

In the event of a trade war, the lobbies' behavior, and therefore the expected trade-war tariffs, are not altered by mean-zero uncertainty (see Equation 2.4): the lobbies will equate the marginal increase in profits to the marginal cost, whether this increase is certain or in expected terms. As this stage is analogous to previous models with a single governmental actor and endogenous political pressure, we can see that the only impact of introducing political uncertainty would be on the legislatures' decision after the realization of θ_n.

However, at the earlier stage of lobbying, uncertainty alters optimizing behavior by both the lobby and the executives. In contrast to the example in the previous section with no uncertainty, consider the case of a very small amount: take θ distributed uniformly on $[-0.01, 0.01]$. Here, the executives find it optimal to set the tariffs once again to completely disengage the lobby and ensure the agreement remains in force. However, the tariff level will be different because instead of making the decision to lobby according to the certainty condition

$$\pi(\tau^n) - e_n - \pi(\tau^a) > e_b,$$

the lobby makes this decision according to Condition 2.8, which can be re-written as

$$\pi(\tau^n) - e_n - \pi(\tau^a) > \left[\frac{\partial b(0, \boldsymbol{\tau^a}, \boldsymbol{\tau^n}(e_n, e_n^*))}{\partial e_b}\right]^{-1}.$$

Although the relationship between the tariff under certainty and a small amount of uncertainty will depend on the structure of the economy and the legislative process, in this case, the tariff can be reduced slightly to 0.105. As we have seen above, with significant uncertainty, the tariff in this example is reduced to 0.078. Interestingly, the

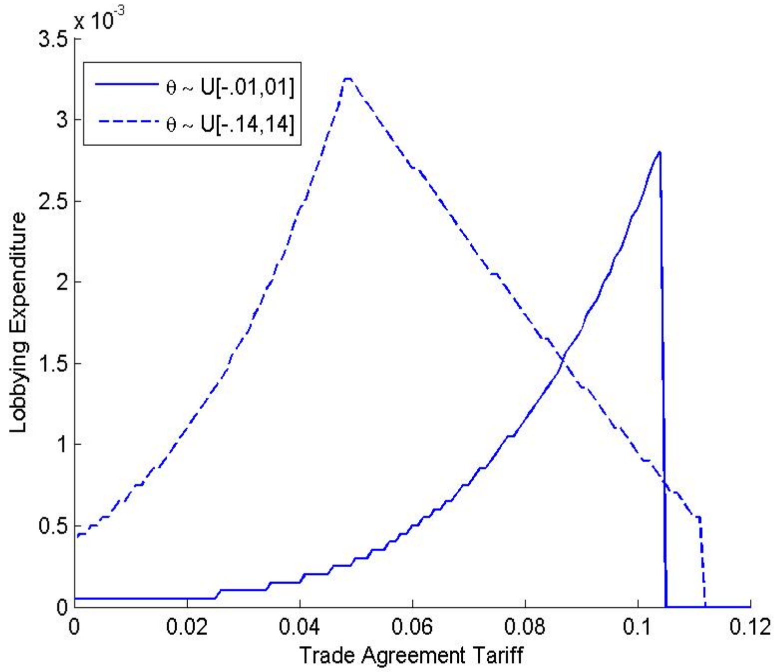

Figure 2.3: Lobby's optimal expenditure function (low and intermediate uncertainty)

optimal tariff does not decrease monotonically between these two values as uncertainty increases.

It is particularly instructive to examine the lobby's reaction function at this very low level of uncertainty as well as the intermediate level when θ is distributed uniformly on $[-0.14, 0.14]$ because it is at the latter that the executives set the highest tariff level of 0.107. For the former, when facing very little uncertainty, the lobby finds it optimal to contribute at a level high enough to ensure the legislature will break the agreement up until $\boldsymbol{\tau^a} = 0.103$ and disengages completely at $\boldsymbol{\tau^a} = 0.105$ (see Figure 2.3). When there is so little uncertainty, the range of tariff levels to which the lobby responds with an intermediate contribution that leaves open the possibility of a break is very small: only those between 0.103 and 0.104.

In contrast, when uncertainty is increased to the interval $[-0.14, 0.14]$, we see that the increasing portion of the reaction function only extends to $\boldsymbol{\tau^a} = 0.048$. After this point, the lobby begins to leave open the possibility of a break and the reaction function begins to decrease in $\boldsymbol{\tau^a}$ as predicted in Result 2. It chooses to disengage completely

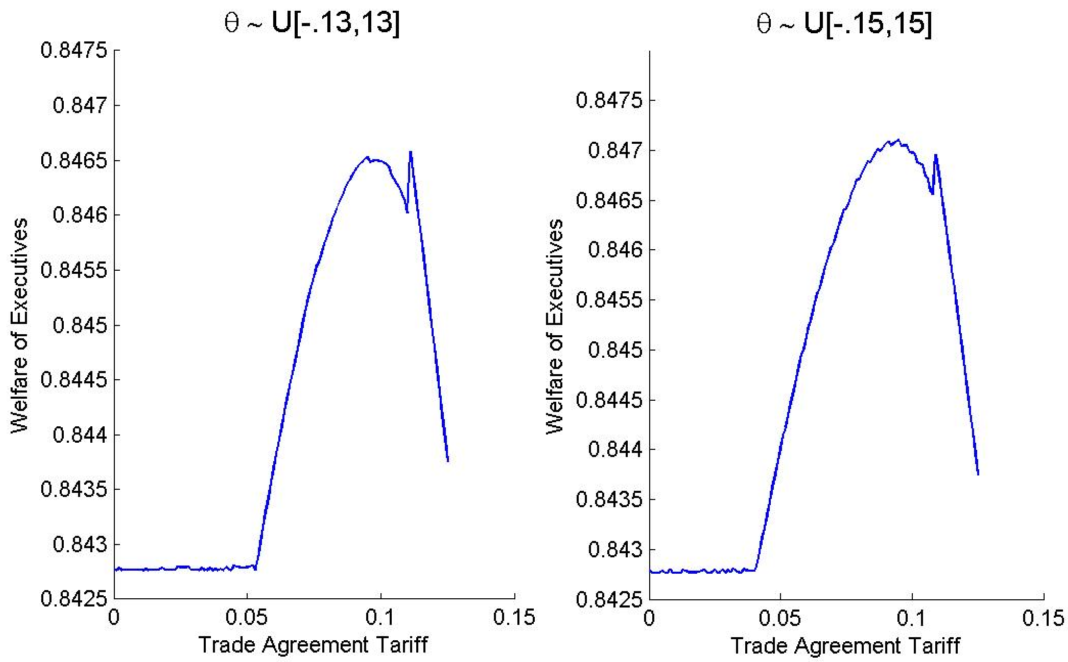

Figure 2.4: Differences in executive welfare when uncertainty increases

at 0.107 and this is the tariff level that the executives find optimal. Interestingly, at intermediate uncertainly levels, both portions of the best response curve are steeper and it turns out that the lobby disengages at slightly lower tariff levels—as low as 0.099 when $\theta \sim U[-0.06, 0.06]$.

The preceding discussion illustrates Part (a) of Result 4: here the executives maximize their welfare by raising tariffs to the point where the import-competing industry ceases to lobby and thus the agreement remains in force for sure. Note, however, that although lobbying expenditures are zero, the potential for lobbying behavior is essential in shaping the trade agreement.

When uncertainty rises above this threshold of $\theta \sim U[-0.14, 0.14]$, the executives' choices are made according to Part (b) of Result 4: that is, they trade off reductions in welfare under the agreement with reductions in the probability that the agreement will be abrogated by the legislature. In terms of the lobby's best response function, the executives now choose the optimal point on the downward-sloping portion of the curve instead of the point where it reaches zero. In terms of the welfare function for the executives, shown in Figure 2.4, they are now choosing the point that maximizes the concave portion of the curve instead of the "spike" that is created when the lobbies disengage and the chance that the agreement will be abrogated is removed altogether.[20]

As uncertainty increases further, the general pattern is that tariffs, and therefore lobbying effort, are reduced, while the break probability increases. However, this pattern is not smooth or absolute, in contrast to the welfare level achieved by the executives, which increases monotonically in the amount of uncertainty present. It is not clear how general a result this is, but it is particularly interesting in contrast to the non-monotone pattern over the lower-range of uncertainty where executive welfare is minimized at the edges of the range that lead to disengagement of the lobby and maximized in the middle of that range.

Several salient points emerge from this example. First, different amounts of uncertainty lead to very different outcomes—both in terms of the tariffs that are set in the trade agreement and the probability of disruption. Second, these outcomes vary in nuanced ways with the underlying behavior of the lobbies and the executives. The provided tariff levels are strongly influenced by lobbying incentives, but not in the straightforward ways predicted by unitary models. Finally, the empirical puzzle surrounding the PFS model can be resolved: in particular, with small enough levels of political uncertainty, this example shows that a welfare-maximizing executive and lobbying effort at zero (and therefore legislature with no endogenous protectionist bias) are consistent with very high tariff levels.

2.5 Extensions

2.5.1 Asymmetric Trade Agreements

The assumption that the legislatures do not have the opportunity to break a trade agreement simultaneously serves to simplify the analysis of the interaction between the

[20]Note that the initial, flat portion of the curve is where the tariff is low enough that the lobby buys a trade war with certainty; we then enter the concave portion where the main results of Section 2.2 are applicable; finally, we see the upward spike where the probability of a trade war is reduced to zero and the executives' welfare declines after that because there is no further reduction in break probability as the tariff increases.

lobbies and the consequences of relaxing it will be discussed in the following section. Maintaining it for now, however, facilitates examination of the assumption that trade agreements must be symmetric. This may be quite realistic in some settings where additional political constraints would make "unfair" arrangements unpalatable. However, it is important to point out that, for some parameter choices, the joint welfare of the executives is in fact maximized by an asymmetric agreement.

The parameterization from the example in Section 2.3 is one such instance. The optimal unrestricted trade agreement is to set tariffs in one country (say home) at 0.062 and those in the other (foreign) at 0. In this case, the equilibrium break probability in home, if the home legislature is afforded the opportunity, is zero, while the foreign legislature will break the agreement with probability 1 if given the chance. With the assumption from the example that each legislature gets the opportunity to repudiate the agreement half the time, this results in the trade agreement being broken with probability 0.50, just less than the 0.505 of the optimal symmetric agreement with both tariffs set at 0.078. Because both the trade agreement tariffs and the expected chance that the agreement will remain in force are lower, expected welfare for the executives is higher.

What is behind the rather surprising result that a lower break probability can be achieved when both tariffs are strictly lower than in the symmetric agreement? Here, the executives relinquish any chance of passage of the trade agreement in the foreign country (in essence hoping that its legislature will not have a chance to act). Once they decide to pursue this course, it is optimal to reduce τ^{*a} to zero. This drastically reduces the home legislature's incentive to break the trade agreement. Because of the large direct impact of τ^{*a} in reducing the break probability, the home tariff is not required to make such a large contribution and can be lowered, which in turn reduces the lobby's incentive to exert effort (see Result 2). This effect is strong enough that the executives can set a tariff level lower than that of the symmetric agreement that is sufficient to disengage the home lobby (see Inequality 2.8).

It is useful to keep in mind that this result occurs in a completely symmetric environment. Recall that, in general, the executives maximize joint welfare of

$$\{o \cdot B(\boldsymbol{\tau^a}) + o^* \cdot B^*(\boldsymbol{\tau^a})\}\, \boldsymbol{W_E}(\boldsymbol{\tau^n}) + \{1 - o \cdot B(\boldsymbol{\tau^a}) - o^* \cdot B^*(\boldsymbol{\tau^a})\}\, \boldsymbol{W_E}(\boldsymbol{\tau^a})$$

If the chance that the legislatures will have the opportunity to break the agreement (o

and o^*) are not equal, or if another significant aspect of the economic or political situation is asymmetric, the importance of this result grows in a potentially substantial way. In particular, if only one country faces a legislative constraint, an asymmetric agreement seems highly likely as in the large body of literature investigating the Schelling conjecture.

2.5.2 Break Opportunities in Both Countries

Although it changes their expected payoffs, extending the model to allow both legislatures the opportunity to break the trade agreement simultaneously does not impact the legislatures' incentives as each decision to break or not break the agreement is independent of the other.

However, the lobbies' incentives are altered because the total probability of a break with which they weight profits changes. Now, in addition to their own effort positively influencing their own legislature to break the agreement, the probability of a break in the trade agreement also increases in the effort of the other country's lobby. As might be expected, the lobbying effort in each country turns out to be a negative function of that in the trading partner; that is, to some extent, the lobbies have an incentive to free ride.

Without making stronger assumptions, it is not possible to say how much the effort of each lobby will be reduced therefore whether total lobbying effort rises or falls. What is clear is that the optimization of the executives becomes significantly more complex as they are now faced with the problem of how to best exploit the free-riding dynamics that occur between the lobbies.

2.6 Conclusion

I have shown that the legislature both breaks trade agreements with a higher probability and sets higher trade war tariffs when lobbying activity increases, while the probability with which it breaks agreements decreases in the domestic trade agreement tariff. Because the lobby decreases its effort in response to higher trade agreement tariffs, the executives face a trade-off between the welfare derived while a trade agreement is in force and the probability with which the agreement is broken.

I have also shown that in a government in which power is separated between branches of government, a less politically-motivated executive can utilize an international trade agreement to reduce the political pressure on the legislative branch and therefore increase the probability that the agreement will remain in force. Thus, in a model with a richer description of government structure, a political-commitment role for trade agreements can arise.

The executives' incentive to raise tariffs in order to reduce lobbying effort, as well as a more realistic legislative structure, helps to explain the empirical finding in the Protection for Sale literature that levels of protection and associated deadweight losses are too high relative to lobbying expenditure given the high estimates for governments' weighting of social welfare. I have shown that both serve to mediate the relationship between preferences for contributions relative to social welfare and the tariffs that are provided. The former has a particularly intuitive facet: the observed lobbying expenditure levels may in fact be low *because* tariffs have been raised sufficiently high to prevent political pressure and the increased risk of a costly trade disruption it engenders.

That lobbying and tariff levels are related in systematic ways to the amount of political uncertainty present suggests interesting avenues for future empirical work. Several directions for future theoretical work also seem potentially fruitful, including removing the assumption of perfect enforceability and supporting cooperation through repeated interaction and generalizing the model to the case of multiple lobbies.

2.7 Appendix

Lemma 6. *When prices are linear in tariffs, the second-order condition for the legislature's problem when setting trade-war tariffs holds for all non-prohibitive tariffs given that the tariff elasticity of supply is less than unity.*

Proof: The first order condition can be rewritten as

$$-D(P_X)\frac{\partial P_X}{\partial \tau} + \gamma(e_n, \theta_n) Q_X(P_X)\frac{\partial P_X}{\partial \tau} + [D(P_X) - Q_X(P_X)] + \tau\left(\frac{\partial D(P_X)}{\partial \tau} - \frac{\partial Q_X(P_X)}{\partial \tau}\right) = 0$$

The required second order condition is that the following is negative:

$$-D(P_X)\frac{\partial^2 P_X}{\partial\tau^2}-\frac{\partial D(P_X)}{\partial\tau}\frac{\partial P_X}{\partial\tau}+\gamma(e_n,\theta_n)Q_X(P_X)\frac{\partial^2 P_X}{\partial\tau^2}+\gamma(e_n,\theta_n)\frac{\partial Q_X(P_X)}{\partial\tau}\frac{\partial P_X}{\partial\tau}$$
$$+2\left[\frac{\partial D(P_X)}{\partial\tau}-\frac{\partial Q_X(P_X)}{\partial\tau}\right]+\tau\left(\frac{\partial^2 D(P_X)}{\partial\tau^2}-\frac{\partial^2 Q_X(P_X)}{\partial\tau^2}\right)$$

$\gamma(e_n,\theta_n)$ is endogenous, so we must find a way to bound it. Because the lobby receives no gains if the tariff is raised beyond its prohibitive level (that is when $D(P_X) = Q_X(P_X)$), it has no incentive to exert effort above the level that induces the $\gamma(e_n,\theta_n)$ that leads to the prohibitive tariff. Setting $D(P_X) = Q_X(P_X)$ in the first order condition, this is

$$\gamma(e_n,\theta_n) = 1-\frac{\tau\left(\frac{\partial D(P_X)}{\partial\tau}-\frac{\partial Q_X(P_X)}{\partial\tau}\right)}{Q_X(P_X)\frac{\partial P_X}{\partial\tau}}$$

If prices are linear in tariffs, the second order terms are zero, and substituting the above expression for $\gamma(e_n,\theta_n)$, we arrive at

$$-\frac{\partial D(P_X)}{\partial\tau}\frac{\partial P_X}{\partial\tau}+\frac{\partial Q_X(P_X)}{\partial\tau}\frac{\partial P_X}{\partial\tau}-\tau\left(\frac{\partial D(P_X)}{\partial\tau}-\frac{\partial Q_X(P_X)}{\partial\tau}\right)\frac{\frac{\partial Q_X(P_X)}{\partial\tau}}{Q_X(P_X)}+$$
$$2\left[\frac{\partial D(P_X)}{\partial\tau}-\frac{\partial Q_X(P_X)}{\partial\tau}\right]$$

Because $P_X = P_X^W + \tau$, $\frac{\partial P_X}{\partial\tau} = \frac{\partial P_X^W}{\partial\tau} + 1$; together with $\frac{\partial P_X^W}{\partial\tau} < 0$, this implies $\frac{\partial P_X}{\partial\tau} < 1$. Therefore the previous expression is bounded above by

$$-\tau\left(\frac{\partial D(P_X)}{\partial\tau}-\frac{\partial Q_X(P_X)}{\partial\tau}\right)\frac{\frac{\partial Q_X(P_X)}{\partial\tau}}{Q_X(P_X)}+\left[\frac{\partial D(P_X)}{\partial\tau}-\frac{\partial Q_X(P_X)}{\partial\tau}\right]$$

Since $\frac{\partial D(P_X)}{\partial\tau}-\frac{\partial Q_X(P_X)}{\partial\tau} < 0$, the second order condition is negative as long as

$$\frac{\tau}{Q_X(P_X)}\frac{\partial Q_X(P_X)}{\partial\tau} < 1.$$

If prices are not linear in the tariff, additional conditions are required. ■

Proof of Lemma 1:

By the implicit function theorem, $\frac{\partial\tau^n}{\partial\gamma} = -\frac{\frac{\partial FOC}{\partial\gamma}}{\frac{\partial FOC}{\partial\tau^n}}$. We have $\frac{\partial FOC}{\partial\gamma} = Q_X(P_X)\frac{\partial P_X}{\partial\tau}$ from the proof of Lemma 6. This is always positive. Lemma 6 also shows that $\frac{\partial FOC}{\partial\tau^n}$ is negative everywhere. ■

Proof of Result 1:

Substituting from Equation 2.1, Equation 2.6 can be re-written as

$$b(e_b, \boldsymbol{\tau^a}, \boldsymbol{\tau^n}) = \Pr[CS_X(\tau^n)+CS_Y(\tau^{*n})+\gamma(e_b,\theta_b)\cdot PS_X(\tau^n)+PS_Y(\tau^{*n})+TR(\tau^n) > CS_X(\tau^a) + CS_Y(\tau^{*a}) + \gamma(e_b,\theta_b)\cdot PS_X(\tau^a) + PS_Y(\tau^{*a}) + TR(\tau^a)|e_b] \quad (2.11)$$

Rearranging, we have

$$b(e_b, \boldsymbol{\tau^a}, \boldsymbol{\tau^n}) = \Pr\Big[\frac{CS_X(\tau^n)+CS_Y(\tau^{*n})+PS_X(\tau^n)+PS_Y(\tau^{*n})+TR(\tau^n)}{PS_X(\tau^a)-PS_X(\tau^n)} - \frac{CS_X(\tau^a)+CS_Y(\tau^{*a})+PS_X(\tau^a)+PS_Y(\tau^{*a})+TR(\tau^a)}{PS_X(\tau^a)+PS_X(\tau^n)} + 1 < \gamma(e_b,\theta_b)|e_b\Big] \quad (2.12)$$

The left side of the inequality in Expression 2.12 does not depend on e_b. But Assumption 1, the right side of the inequality is increasing and concave in e_b. Thus $b(e_b w, \boldsymbol{\tau^a}, \boldsymbol{\tau^n})$ is increasing and concave in e_b. ■

Proof of Lemma 2:

It must be shown that the left hand side of the inequality in Expression 2.12 is decreasing in τ^{*a}. The derivative of this quantity with respect to τ^{*a} is

$$\frac{-\left(PS_X(\tau^a) - PS_X(\tau^n)\right)\left(\frac{\partial CS_Y(\tau^{*a})}{\partial \tau^{*a}} + \frac{\partial PS_Y(\tau^{*a})}{\partial \tau^{*a}}\right)}{\left(PS_X(\tau^a) - PS_X(\tau^n)\right)^2} = \frac{\frac{\partial CS_Y(\tau^{*a})}{\partial \tau^{*a}} + \frac{\partial PS_Y(\tau^{*a})}{\partial \tau^{*a}}}{PS_X(\tau^n) - PS_X(\tau^a)}. \quad (2.13)$$

The reduction in price under the agreement causes a decrease in producer surplus and an increase in consumer surplus. Because Y is the export good, the decrease in producer surplus is larger than the increase in consumer surplus, making the numerator, and thus the entire quantity (since producer surplus is increasing in τ), negative. ■

Proof of Lemma 3:

Using the logic of the proof of Lemma 2, the effect on the break probability is determined by the sign of the derivative of the left hand side of the inequality in Expression 2.12 with respect to τ^a; to show that the break probability is decreasing in τ^a, I must demonstrate that this derivative is positive. Labeling the numerator of that expression $[W(\boldsymbol{\tau^n}) - W(\boldsymbol{\tau^a})]$ (for the change in social welfare), this derivative can be written

$$\frac{(PS_X(\tau^n)-PS_X(\tau^a))\left(\frac{\partial CS_X(\tau^a)}{\partial\tau^a}+\frac{\partial PS_X(\tau^a)}{\partial\tau^a}+\frac{\partial TR(\tau^a)}{\partial\tau^a}\right)-[W(\boldsymbol{\tau^n})-W(\boldsymbol{\tau^a})]\frac{\partial PS_X(\tau^a)}{\partial\tau^a}}{(PS_X(\tau^a)-PS_X(\tau^n))^2}. \tag{2.14}$$

$(PS_X(\tau^n) - PS_X(\tau^a))$ is always positive by Assumption 2. Because the optimal unilateral tariff for large welfare-maximizing governments is positive (call it τ^O),

$$\left(\frac{\partial CS_X(\tau^a)}{\partial\tau^a}+\frac{\partial PS_X(\tau^a)}{\partial\tau^a}+\frac{\partial TR(\tau^a)}{\partial\tau^a}\right)$$

is increasing up to τ^O and decreasing above it. Thus the first summand is increasing up until τ^O and decreasing thereafter.

Because total social welfare is maximized at $\tau^a = \tau^{*a} = 0$, $W(\boldsymbol{\tau^n}) - W(\boldsymbol{\tau^a})$ is always negative, whereas producer surplus is increasing in τ^a, so the second summand is positive everywhere. With a positive denominator, we thus have that the derivative is positive on $[0, \tau^O]$.

It is also positive over the remaining (τ^O, τ^n). To see this, notice that one can add

$$\left(\tilde{\Gamma}-1\right)\frac{\partial PS_X(\tau^a)}{\partial\tau^a}\left(PS_X(\tau^n)-PS_X(\tau^a)\right)$$

to the first summand and subtract it from the second. For any particular value of $\tilde{\tau}^a$, one can choose the $\tilde{\Gamma}$ weight that would make $\tilde{\tau}^a$ the preferred unilateral tariff; this makes the derivative in the first summand zero. Having subtracted the same quantity from the second summand modifies the welfare difference in the second summand to be maximized at $\tilde{\tau}^a$ so that this term is always negative, thus ensuring the result. ■

Proof of Result 2:

Proof is via the Implicit Function Theorem using the lobby's first order condition, Equation 2.7, referred to here as FOC_L.

$$\frac{\partial e_b}{\partial\tau^a} = -\frac{\frac{\partial FOC_L}{\partial\tau^a}}{\frac{\partial FOC_L}{\partial e_b}} = \frac{\frac{\partial b}{\partial e_b}\frac{\partial\pi(\boldsymbol{\tau^a})}{\partial\tau^a} - \frac{\partial^2 b}{\partial e_b\partial\tau^a}\left[\pi(\tau^n)-\pi(\tau^a)\right]}{\frac{\partial^2 b}{\partial {e_b}^2}\left[\pi(\tau^n)-\pi(\tau^a)\right]}$$

From the conclusion of the Proof of Result 1, we can easily take the cross-derivative to see that $\frac{\partial^2 b}{\partial e_b\partial\tau^a}$ is zero since $\gamma(e_b,\theta_b)$ does not depend on τ^a. $\frac{\partial b}{\partial e_b}$ is positive and $\frac{\partial^2 b}{\partial {e_b}^2}$ negative by Result 1 while $\frac{\partial\pi(\boldsymbol{\tau^a})}{\partial\tau^a}$ is positive by construction.

Because $\pi(\tau)$ is increasing everywhere, $[\pi(\tau^n) - e_n - \pi(\tau^a)]$ is positive for all but very large values of τ^a, that is for all τ^a such that $\pi(\tau^n) - e_n > \pi(\tau^a)$. For these values, $\frac{\partial e_b}{\partial \tau^a} < 0$. When τ^a rises above this level, it is no longer in the lobby's interest to ask to have the agreement broken so $e_b = 0$ and $\frac{\partial e_b}{\partial \tau^a} = 0$ Thus $\frac{\partial e_b}{\partial \tau^a} \leq 0$. ■

Proof of Corollary 1:

As in the proof of Result 2, I employ the Implicit Function Theorem on the lobby's first order condition, Equation 2.7, denoted FOC_L.

$$\frac{\partial e_b}{\partial \tau^{*a}} = -\frac{\frac{\partial FOC_L}{\partial \tau^{*a}}}{\frac{\partial FOC_L}{\partial e_b}} = \frac{\frac{\partial b}{\partial e_b}\frac{\partial \pi(\boldsymbol{\tau^a})}{\partial \tau^{*a}} - \frac{\partial^2 b}{\partial e_b \partial \tau^{*a}}\left[\pi(\tau^n) - e_n - \pi(\tau^a)\right]}{\frac{\partial^2 b}{\partial {e_b}^2}\left[\pi(\tau^n) - e_n - \pi(\tau^a)\right]}$$

From the conclusion of the Proof of Result 1, we can take the cross-derivative to see that $\frac{\partial^2 b}{\partial e_b \partial \tau^{*a}}$ is zero since $\gamma(e_b, \theta_b)$ does not depend on τ^{*a}. $\frac{\partial \pi(\boldsymbol{\tau^a})}{\partial \tau^{*a}}$ is also zero: because of the separability between the sectors, profits in the import-competing sector do not depend on τ^{*a}. Thus $\frac{\partial e_b}{\partial \tau^{*a}} = 0$. ■

Proof of Lemma 4:

Again, I want to show how the inequality in Expression 2.12 changes, now with respect to both τ^a and τ^{*a}, so I add the derivatives in Expressions 2.13 and 2.14 to get

$$\frac{\left(PS_X(\tau^n) - PS_X(\tau^a)\right)\left(\frac{\partial W_X(\boldsymbol{\tau^a})}{\partial \tau^a} + \frac{\partial W_X(\boldsymbol{\tau^a})}{\partial \tau^{*a}}\right) - \left[W(\boldsymbol{\tau^n}) - W(\boldsymbol{\tau^a})\right]\frac{\partial PS_X(\tau^a)}{\partial \tau^a}}{\left(PS_X(\tau^n) - PS_X(\tau^a)\right)^2}.$$

where $\frac{\partial W_X(\boldsymbol{\tau^a})}{\partial \tau^a} + \frac{\partial W_X(\boldsymbol{\tau^a})}{\partial \tau^{*a}} = \frac{\partial CS_X(\tau^a)}{\partial \tau^a} + \frac{\partial PS_X(\tau^a)}{\partial \tau^a} + \frac{\partial TR(\tau^a)}{\partial \tau^a} + \frac{\partial CS_Y(\tau^{*a})}{\partial \tau^{*a}} + \frac{\partial PS_Y(\tau^{*a})}{\partial \tau^{*a}}$ is the total derivative of social welfare. Since social welfare is maximized at $\boldsymbol{\tau^a} = (0,0)$,[21] this is negative $\forall \boldsymbol{\tau} \in (\mathbf{0}, \boldsymbol{\tau^n}]$; note that it is 0 at $\mathbf{0}$ and vanishingly small for very small tariffs.

Thus the first summand in the numerator is zero at $\boldsymbol{\tau^a} = \mathbf{0}$ and increasingly negative as $\boldsymbol{\tau^a}$ increases. The second summand is positive everywhere because social welfare, W, is lowest at $\boldsymbol{\tau^n}$ and producer surplus is increasing everywhere. Thus the numerator is positive at $\mathbf{0}$ and at least for very small $\boldsymbol{\tau^a}$

[21]Note, this is identical to the result that joint social welfare is maximized at zero tariffs because of symmetry.

It is also positive for all other values of $\boldsymbol{\tau^a}$ strictly below $\boldsymbol{\tau^n}$. Just as in the proof of Lemma 3, one one can add

$$\left(\tilde{\Gamma} - 1\right) \frac{\partial PS_X(\tau^a)}{\partial \tau^a} \left(PS_X(\tau^n) - PS_X(\tau^a)\right)$$

to the first summand and subtract it from the second. For any particular value of $\tilde{\boldsymbol{\tau^a}}$, one can choose the $\tilde{\Gamma}$ weight that would make $\tilde{\boldsymbol{\tau^a}}$ the politically optimal tariff; this makes the derivative in the first summand zero. Having subtracted the same quantity from the second summand modifies the welfare difference in the second summand to be maximized at $\tilde{\boldsymbol{\tau^a}}$ so that this term is always negative, thus ensuring the result.

Because the denominator is positive, the entire expression is positive for all $\boldsymbol{\tau^a} < \boldsymbol{\tau^n}$. ■

Proof of Result 3:

$$\frac{\partial B(\boldsymbol{\tau^a})}{\partial \tau^a} + \frac{\partial B(\boldsymbol{\tau^a})}{\partial \tau^{*a}} = \frac{\partial b}{\partial e_b}\frac{\partial e_b}{\partial \tau^a} + \frac{\partial b}{\partial e_b}\frac{\partial e_b}{\partial \tau^{*a}} + \frac{\partial b}{\partial \tau^a} + \frac{\partial b}{\partial \tau^{*a}}$$

The first summand is negative by Results 1 and 2. As shown in Corollary 1, the second summand is zero because $\frac{\partial e_b}{\partial \tau^{*a}}$ is zero. Taken together, the final two summands are negative by Lemma 4. Thus the entire expression is negative. ■

Conditions for Interior Solution to Executives' Problem

The first order condition for maximizing joint surplus with respect to τ^a and τ^{*a} when the two are constrained to be equal is

$$\{1 - o \cdot B(\boldsymbol{\tau^a}) - o^* \cdot B^*(\boldsymbol{\tau^a})\} \left\{\frac{\partial \boldsymbol{W_E}(\boldsymbol{\tau^a})}{\partial \tau^a} + \frac{\partial \boldsymbol{W_E}(\boldsymbol{\tau^a})}{\partial \tau^{*a}}\right\}$$
$$+\left\{o \cdot \left[\frac{\partial B(\boldsymbol{\tau^a})}{\partial \tau^a} + \frac{\partial B(\boldsymbol{\tau^a})}{\partial \tau^{*a}}\right] + o^* \cdot \left[\frac{\partial B^*(\boldsymbol{\tau^a})}{\partial \tau^a} + \frac{\partial B^*(\boldsymbol{\tau^a})}{\partial \tau^{*a}}\right]\right\} [\boldsymbol{W_E}(\boldsymbol{\tau^n}) - \boldsymbol{W_E}(\boldsymbol{\tau^a})] = 0$$

Because there is no benefit to setting $\boldsymbol{\tau^a}$ below the executives' preferred level, which I will denote $\boldsymbol{\tau^E}$, I will take the choice space to be $[\boldsymbol{\tau^E}, \boldsymbol{\tau^n}]$. Note that for $\gamma_E = 1$, $\boldsymbol{\tau^E} = 0$.

To demonstrate that the executives do not choose $\boldsymbol{\tau^a} = \boldsymbol{\tau^E}$, I must show that the left side of the above equation is positive at $\boldsymbol{\tau^a} = \boldsymbol{\tau^E}$. Assumption 2 and Result 3 combined with symmetry ensure that the first term of the second summand is negative. That executive welfare is maximized at $\boldsymbol{\tau^E}$ ensures that the term multiplying it is also negative as well as that $\frac{\partial \boldsymbol{W_E}(\boldsymbol{\tau^E})}{\partial \tau^a} + \frac{\partial \boldsymbol{W_E}(\boldsymbol{\tau^E})}{\partial \tau^{*a}}$ is zero. Therefore the derivative of joint executive welfare is positive at $\boldsymbol{\tau^E}$. ■

2.8 References

Aidt, Toke and Martin Gassebner (2010), "Do Autocratic States Trade Less?," *The World Bank Economic Review*, 24, 38-76.

Ansolabehere, Stephen, John M. de Figueiredo and James M. Snyder Jr. (2003), "Why Is There so Little Money in U.S. Politics?," *Journal of Economic Perspectives*, 17, 105-130.

Bagwell, Kyle and Robert Staiger (2005), "Enforcement, Private Political Pressure, and the General Agreement on Tariffs and Trade/World Trade Organization Escape Clause," *Journal of Legal Studies*, 34, 471-513.

Baldwin, Robert E., (1987), "Politically realistic objective functions and trade policy: PROFs and tariffs," *Economic Letters*, 24, 287290.

Beshkar, Mostafa (2010), "Trade skirmishes safeguards: A theory of the WTO dispute settlement process," *Journal of International Economics*, 82, 35-48.

Bombardini, Matilde (2008), "Firm Heterogeneity and Lobby Participation," *Journal of International Economics*, 75, 329-348.

Bombardini, Matilde and Francesco Trebbi (2009), "Competition and Political Organization: Together or Alone in Lobbying for Trade Policy?" NBER Working Paper 14771.

Dai, Xinyuan (2006), "Dyadic Myth and Monadic Advantage: Conceptualizing the Effect of Democratic Constraints on Trade," *Journal of Theoretical Politics*, 18, 267-297.

Destler, I.M. (2005), "American Trade Politics," Institute for International Economics, 4th ed.

Feenstra, Robert (1992), "How Costly is Protectionism?" *Journal of Economic Perspectives*, 6, 159-178.

Feenstra, Robert and Tracy Lewis (1991), "Negotiated Trade Restrictions with Private Political Pressure," *Quarterly Journal of Economics*, 106, 1287-1307.

Gawande, Kishore and Usree Bandyopadhyay (2000), "Is Protection for Sale? Evidence on the Grossman-Helpman Theory of Endogenous Protection," *The Review of Economics and Statistics*, 82, 139-152.

Gawande, Kishore and Bernard Hoekman (2006), "Lobbying and Agricultural Trade Policy in the United States," *International Organization*, 60, 527-561.

Gawande, Kishore and Christopher Magee (2011), "Free Riding on Protection for Sale," *International Studies Quarterly*.

Gawande, Kishore, Pravin Krishna and Michael Robbins (2006), "Foreign Lobbies and U.S. Trade Policy," *Review of Economics and Statistics*, 88, 563-571.

Gawande, Kishore, Pravin Krishna and Marcelo Olarreaga (2005), "Lobbying Competition Over Trade Policy," NBER Working Paper 11371.

Goldberg, Pinelopi and Giovanni Maggi (1999), "Protection for Sale: An Empirical Investigation," *American Economic Review*, 89, 1135-1155.

Grossman, Gene and Elhanan Helpman (1994), "Protection for Sale," *The American Economic Review*, 84, 833-850.

Grossman, Gene and Elhanan Helpman (1995), "Trade Wars and Trade Talks," *The Journal of Political Economy*, 103, 675-708.

Henisz, Witold J. and Edward D. Mansfield (2006), "Votes and Vetoes: The Political Determinants of Commercial Openness," *International Studies Quarterly*, 50, 189-211.

Iida, Keisuke (1996), "Involuntary Defection in Two-Level Games," *Public Choice*, 89, 283-303.

Imai, Susumu, Hajime Katayama and Kala Krishna (2009), "Is protection really for sale? A survey and directions for future research," *International Review of Economics and Finance*, 18, 181-191.

Kibris, Arzu (2010), "Uncertainty and Ratification Failure," *Public Choice*, available at: http://dx.doi.org/10.1007/s11127-010-9710-x.

Klimenko, Mikhail, Garey Ramey and Joel Watson (2008), "Recurrent Trade Agreements and the Value of External Enforcement," *Journal of International Economics*, 74, 475-499.

Kono, Daniel Y. (2006), "Optimal Obfuscation: Democracy and Trade Policy Transparency," *American Political Science Review*, 100, 369-384.

Laver, Michael and Kenneth A. Shepsle (1991), "Divided Government: America is Not 'Exceptional'," *Governance: An International Journal of Policy and Administration*, 4, 250-269.

Le Breton, Michel and François Salanie (2003), "Lobbying under political uncertainty," *Journal of Public Economics*, 87, 2589-2610.

Le Breton, Michel and Vera Zaporozhets (2007), "Legislative Lobbying under Political Uncertainty," Available at SSRN: http://ssrn.com/abstract=1024686.

Lohmann, Susanne and Sharyn OHalloran (1994), "Divided Government and US Trade Policy," *International Organization*, 48, 595-632.

Maggi, Giovanni and Robert Staiger (2011), "The Role of Dispute Settlement Procedures in International Trade Agreements," *Quarterly Journal of Economics*, 126, 475-515.

Mansfield, Edward D., Helen Milner and B. Peter Rosendorff (2000), "Free to Trade: Democracies, Autocracies, and International Trade," *The American Political Science Review*, 94, 305-321.

Milner, Helen and Keiko Kubota (2005), "Why the Move to Free Trade? Democracy and Trade Policy in the Developing Countries," *International Organization*, 59, 107-143.

Milner, Helen and B. Peter Rosendorff (1997), "Democratic Politics and International trade Negotiations: Elections and Divided Government as Constraints on Trade Liberalization," *Journal of Conflict Resolution*, 41, 117-147.

Milner, Helen and B. Peter Rosendorff (2001), "The Optimal Design of International Trade Institutions: Uncertainty and Escape," *International Organization*, 55, 829-857.

Mitra, Devashish, Dimitrios D. Thomakos and Mehmet A. Ulubaşoğlu (2002), "'Protection for Sale' in a Developing Country: Democracy vs. Dictatorship," *Review of Economics and Statistics*, 84, 497-508.

Mitra, Devashish, Dimitrios D. Thomakos and Mehmet A. Ulubaşoğlu (2006), "Can We Obtain Realistic Parameter Estimates for the Protection for Sale Model?" *Canadian Journal of Economics*, 39, 187-210.

Morrow, James D., Randolph M. Siverson and Tressa E. Tabares (1999), 'Correction to: "The Political Determinants of International Trade,"' *American Political Science Review*, 94, 931-933.

McCalman, Phillip (2004), "Protection for Sale and Trade Liberalization: an Empirical Investigation," *Review of International Economics*, 12, 81-4.

Paltseva, Elena (2011), "Protection for Sale to Oligopolists," available at: http://web.econ.ku.dk/okoep/Paltseva_P4S2O_June2011_1.pdf.

Song, Yeongkwan (2008), "Protection for Sale: Agenda-Setting and Ratification in the Presence of Lobbying," *Korea Institute for International Trade Policy Working Paper Series* Vol. 2008-22, August 2008.

Tarar, Ahmer (2001), "International Bargaining with Two-Sided Domestic Constraints," *Journal of Conflict Resolution*, 45, 320-340.

Ward, Michael D. and Peter D. Hoff (2007), "Persistent Patterns of International Commerce," *Journal of Peace Research*, 44, 157-175.

Chapter 3

Self-enforcing trade agreements, dispute settlement and separation of powers

Abstract

If external enforcement of international trade agreements is not available in an environment in which one branch of government retains final authority over tariff levels (call it the legislature) while delegating the task of trade-agreement formation to another (call it the executive), the inability of actors to make commitments affects the design of trade agreements in several ways. Executives must not only take into account the legislatures' lobbying-driven propensity to revoke delegation and break the agreement, but also that their agreements must be robust to the executives' own incentives to renegotiate out of any punishment scheme. The design of the dispute settlement mechanism that makes the optimal punishment incentive compatible must balance two, often-conflicting, objectives: longer punishment periods help to enforce cooperation by increasing the costs of defecting from the agreement, but because the lobbies prefer the punishment outcome, this also incentivizes lobbying effort and with it the political pressure to break the agreement. Thus the model generates new predictions for the optimal design of mechanisms for resolving the disputes that arise in the course of trade-agreement relationships.

KEYWORDS: Trade Policy, International Agreements, Lobbying, Political Economy, Structure of Government, Repeated Games, Renegotiation, Dispute Settlement Institution

In the absence of strong external enforcement mechanisms for international trade agreements, we generally assume that cooperation is enforced by promises of future cooperation, or, alternatively, promises of future punishments for exploitative behavior.[1] Klimenko, Ramey and Watson (2008) show that the typical grim trigger punishments are not useful for supporting cooperation when renegotiation is possible. They propose a notion of *recurrent agreement* that takes into account the possibility of renegotiation via one appealing solution: a dispute settlement institution (DSI) loosely patterned on the Dispute Settlement Body of the World Trade Organization that helps trading partners to credibly condition their negotiations on the state of their relationship and avoid the problems created by renegotiation.

Relative to the existing literature, this paper incorporates a separation-of-powers policy-making process as in Buzard (2012) with endogenous lobbying along the lines of Grossman and Helpman (1994 and 1995) into such a repeated-game setting with a DSI to take account of the threat to cooperation posed by renegotiation. The structure is therefore similar to that of Bagwell and Staiger (2005) with endogenously-determined political economy weights and power over the policy-making process modeled as shared between executive and legislative branches of the government as in Milner and Rosendorff (1997) and Song (2008). The recurrent agreement approach following Klimenko, Ramey and Watson 2008 (hereafter KRW) complements contributions by Cotter and Mitchell (1997), Ludema (2001) and Beshkar (2011) that study "renegotiation-proof" trade agreements, as well as Maggi and Staiger (2011) who examine optimal trade agreement design in the presence of costly renegotiation.

Here, welfare-maximizing executives use the trade-agreement as a kind of political commitment device: by setting tariffs to optimally reduce lobbying incentives, they help the legislatures resist political pressure they would otherwise face to break the agreement.[2] Given that all actors have perfect information about the effect of lobbying

[1] Repeated non-cooperative game models of trade agreements have been considered by McMillan (1986, 1989), Dixit (1987), Bagwell and Staiger (1990, 1997a,b, 2002), Riezman (1991), Kovenock and Thursby (1992), Maggi (1999) and Ederington (2001) and others.

[2] The commonly-made assumption that the executive is less protectionist than the legislature is a special case of the finding that susceptibility to special interests generally declines with the size of one's constituency. One simple illustration from the realm of trade policy is the following: a legislator whose district has a large concentration of a particular industry does not take into account the impact of tariffs on the welfare of consumers in other districts, while the executive, whose constituency encompasses the

effort on the outcome of political process, the executives maximize social welfare by choosing the lowest tariffs that make it unattractive for the lobbies to work to provoking a trade dispute. Thus the problem with the lobby adds an extra constraint to the standard problem. The standard constraint on the key repeated-game player, which here is the legislature, is loosened by increasing the punishment length because defections become relatively more unattractive. However, this new constraint due to the presence of lobbying becomes tighter as the punishment becomes more severe because the lobby *prefers* punishment periods. Because the tariffs during punishment, and thus its profits, are higher than during a cooperative period, the lobby has increased incentive to exert effort as the punishment lengthens.

The optimal punishment length must balance these two competing forces. Where the balance falls depends in large part on how influential the lobby is in the legislative process. If the lobby has very little power, the optimal punishment converges to that of the model without a lobby: longer punishments are better because the key constraint is the legislature's. As the lobby becomes stronger, the optimal punishment becomes shorter because the lobby's incentive becomes more important.

I begin by describing the stage game in detail. Section 3.2 sets out the dispute settlement institution and the set-up of the repeated game. I then describe the structure of optimal trade agreements in Section 3.3. Section 3.4 explores the forces shaping optimal dispute resolution procedures and Section 3.5 concludes.

3.1 Stage Game

I employ a partial-equilibrium model with two countries: home (no asterisk) and foreign (asterisk). The countries trade two goods, X and Y, where P_i denotes the home price of good $i \in \{X, Y\}$ and P_i^* denotes the foreign price of good i. In each country, the demand functions are taken to be identical for both goods, respectively $D(P_i)$ in home and $D(P_i^*)$ in foreign and are assumed strictly decreasing and twice continuously differentiable.

The supply functions for good X are $Q_X(P_X)$ and $Q_X^*(P_X^*)$ and are assumed

whole country, will internalize these diffuse consumption effects. For a detailed argument, see Lohmann and O'Halloran (1994).

strictly increasing and twice continuously differentiable for all prices that elicit positive supply. I also assume $Q^*_X(P_X) > Q_X(P_X)$ for any such P_X so that the home country is a net importer of good X. The production structure for good Y is taken to be symmetric, with both demand and supply such that the economy is separable in goods X and Y. It is assumed that the production of each good requires the possession of a sector-specific factor that is available in inelastic supply and is non-tradable so that the income of owners of the specific factors is tied to the price of the good in whose production their factor is used.

For simplicity, I assume each government's only trade policy instrument is a specific tariff on its import-competing good: the home country levies a tariff τ on good X while the foreign country applies a tariff τ^* to good Y. Local prices are then $P_X = P^W_X + \tau$, $P^*_X = P^W_X$, $P_Y = P^W_Y$ and $P^*_Y = P^W_Y + \tau^*$ where a W superscript indicates world prices and equilibrium prices are determined by the market clearing conditions

$$M_X(P_X) = D(P_X) - Q_X(P_X) = Q^*_X(P^*_X) - D(P^*_X) = E^*_X(P^*_X)$$

$$E_Y(P_Y) = Q_Y(P_Y) - D(P_Y) = D(P^*_Y) - Q^*_Y(P_Y) = M^*_Y(P^*_Y)$$

where M_X are home-county imports and E^*_X are foreign exports of good X and E_Y are home-county exports and M^*_Y are foreign imports of good Y.

It follows that P^W_X and P^W_Y are decreasing in τ and τ^* respectively, while P_X and P^*_Y are increasing in the respective domestic tariff. This gives rise to a standard terms-of-trade externality. As profits and producer surplus (identical in this model) in a sector are increasing in the price of its good, profits in the import-competing sector are also increasing in the domestic tariff. This economic fact, combined with the assumptions on specific factor ownership, is what motivates political activity.

I next describe the politically-relevant actors. In order to focus attention on protectionist political forces, I assume that only the import-competing industry in each country is politically-organized and able to lobby and that it is represented by a single lobbying organization. Each country's government is composed of two branches: an executive who can conclude trade agreements and a legislature that has final say on trade policy. In summary, the political process is modeled as involving three players in each country: the lobby, the executive, and the legislature.

The stage-game timing is as follows. First, the executives set trade policy cooperatively in an international agreement. In the context of the repeated game, this can be construed as concluding an agreement in the first period and then potentially renegotiating at the beginning of each subsequent period, or as forming a new agreement each period. After the trade agreement is concluded in each period, the lobbies attempt to persuade the legislators in their respective countries to break the trade agreement. Next, the legislatures decide whether to abide by the agreement or to provoke a trade war. In the event that the trade agreement does not remain in force, there is a final stage of lobbying and voting to set the trade-war tariffs. Once all political decisions are taken, producers and consumers make their decisions.

For the main analysis, I will assume complete information, so the appropriate solution concept is subgame perfect Nash equilibrium.[3] As this game is solved by backward induction, it is intuitive to start by describing the incentives of the legislatures, whose decisions I model as being taken by a median legislator. As the economy is fully separable and the economic and political structures are symmetric, I focus here on the home country and the X-sector. The details are analogous for Y and foreign.

The per-period welfare function of the home legislature is

$$W_{ML} = CS_X(\tau) + CS_Y(\tau^*) + \gamma(e) \cdot PS_X(\tau) + PS_Y(\tau^*) + TR(\tau) \qquad (3.1)$$

where CS is consumer surplus, PS is producer surplus, $\gamma(e)$ is the weight placed on producer surplus (profits) in the import-competing industry, e is lobbying effort, and TR is tariff revenue. Here, the weight the median legislator places on the profits of the import-competing industry, $\gamma(e)$ is affected by the level of lobbying effort.[4]

Assumption 3. $\gamma(e)$ *is continuously differentiable, strictly increasing and concave in* e.

Assumption 3 formalizes the intuition that the legislature favors the import-competing industry more the higher is its lobbying effort, but that there are diminishing

[3] In an extension to the case of political uncertainty, information about that uncertainty will be symmetric and so subgame perfect Nash equilibrium will remain the appropriate solution concept.

[4] The standard PFS modeling would specify $W_{ML} = C + aW$, but as will be seen when we come to the preferences of the executive, this is not sufficiently general for the purposes of this model. Although complex, an isomorphism can be made between the two forms in a special case as discussed in Buzard 2012.

returns to lobbying activity.

Lobbying affects only the weight the legislature places on the profits of the import-competing industry. These profits are higher in a trade war than under a trade agreement, so given Assumption 3, γ is increasing in lobbying effort, implying that the legislature becomes more favorably inclined toward the high trade-war tariff and associated profits as lobbying increases and therefore more likely to break the trade agreement.

Given the legislature's preferences, the home lobby chooses its lobbying effort (e_b to influence the break decision and e_{tw} to influence the trade war tariff) to maximize the welfare function:

$$U_L = [\pi(\tau^{tw}) - e_{tw}]\, \mathbb{1}[\text{Trade War}] + \pi(\tau^{a})\mathbb{1}[\text{TradeAgreement}] - e_b \tag{3.2}$$

where $\pi(\cdot)$ is the current-period profit and τ^a (τ^{tw}) is the home country's tariff on the import good under a trade agreement (war). I use the convention throughout of representing a vector of tariffs for both countries (τ, τ^*) as a single bold $\boldsymbol{\tau}$.

I assume the lobby's contribution is not observable to the foreign legislature. The implication is that the lobby can directly influence only the home legislature, and so the influence of one country's lobby on the other country's legislature occurs only through the tariffs selected.[5]

In the first stage, the executives choose the trade agreement tariffs $\boldsymbol{\tau^a} = (\tau^a, \tau^{*a})$ via a negotiating process that I assume to be efficient. This process therefore maximizes the joint payoffs of the trade agreement:[6]

$$\boldsymbol{W_E}(\boldsymbol{\tau^a}) = W_E(\boldsymbol{\tau^a}) + W_E^*(\boldsymbol{\tau^a}) \tag{3.3}$$

I model the executives' choice via the Nash bargaining solution where the disagreement point is the executives' welfare resulting from the Nash equilibrium in the non-cooperative game (i.e. in the absence of a trade agreement) between the legislatures.

The executives are assumed, for simplicity, to be social-welfare maximizers.

[5] cfr. Grossman and Helpman 1995, page 685.

[6] In the extension to the case of political uncertainty, the joint payoffs must take into account the possibility that the trade agreement will be broken. In the case of certainty, agreement will always be maintained on the equilibrium path and so this specification is sufficient.

Therefore the home executive's welfare is specified as follows:

$$W_E = CS_X(\tau) + CS_Y(\tau^*) + PS_X(\tau) + PS_Y(\tau^*) + TR(\tau)$$

Note that this is identical to the welfare function for the legislature aside from the weight on the profits of the import industry, which is not a function of lobbying effort. This assumption does *not* require that the executives are not lobbied; only that their preferences are not directly altered in a significant way by lobbying over trade—that they do not sell protection in order to finance their re-election campaigns. In the case of the post-war United States, where the Congress has consistently been significantly more protectionist than the President, this seems to reasonably reflect the political reality. For trade policy, where there are concentrated benefits but harm is diffuse, there are good reasons for this to be the case. Because the President has the largest constituency possible, delegating authority to the executive branch may simply be a mechanism for "concentrating" the benefits since consumers seem unable to overcome the free-riding problem. In fact, a strong argument can be made that power over trade policy has been delegated to the executive branch precisely *because* it is less susceptible to the influence of special interests (Destler 2005).

Therefore, in line with both the theoretical and empirical literature, I will assume that $\gamma(e) \geq 1$ for all e. That is, even for the least favorable outcome of the lobbying process, the legislature will be at least slightly more protectionist than the executive.

Assumption 4. $\gamma(e) \geq 1 \ \forall e.$

Assumption 4 ensures that $\tau^a < \tau^{tw}$, and more generally, that the legislature's incentives are more closely aligned with the lobby's than are those of the executive. This is not essential but simplifies the analysis and matches well the empirical findings that politicians with larger constituencies are less sensitive to special interests (See Destler 2005 and footnote 2 above).

Although the political process here matches most closely that of the United States in the post-war era, I believe the model or one of its extensions is applicable for a broad range of countries for which authority over the formation and maintenance of trade policy is diffuse and subject to political pressure either at home or in a trading partner.[7]

3.2 Repeated Game

3.2.1 Dispute Settlement Institution

Following KRW, I will assume that the countries submit themselves to an external Dispute Settlement Institution (DSI) for the purposes of overcoming the renegotiation problem: that is, the incentive to renegotiate out of punishment phases that destroys the ability of the punishments to enforce cooperation. One way to (informally) make adherence to the DSI incentive compatible is to imagine that many trading partners use the DSI and that all will punish a country who deviates in any bilateral agreement.

The DSI is assumed to keep records of the negotiated agreements, complaints, and violations, and to settle disputes when agreements are violated. The simple DSI employed here conditions the interaction of the countries in the following manner.[8] The DSI keeps records in terms of two possible states of the trade relationship, "cooperative" and "dispute." At the start of any period, it is assumed that either there is no dispute pending, or else the DSI is in the process of resolving a dispute triggered by a violation in some prior period. I refer to the former situation as the "cooperative state," or state C. If a dispute is pending, then the period begins in the "dispute state," or state D. When a tariff agreement is violated, the DSI switches the state from C to D, and a dispute settlement process (DSP) begins, as described below. When settlement is achieved, the DSI switches the state from D back to C.

Importantly, the DSI cannot be directly manipulated by the countries involved in a dispute, so countries continue to negotiate agreements and choose tariffs as before, except their negotiation can be conditioned on the DSI's state. Therefore, the negotiation problem that countries face following a dispute history may be different than the negotiation problem they face following a cooperative history.

[7] In particular, the binary decision by the legislature about whether to abide by or break the trade agreement is modeled on the "Fast Track Authority" that the U.S. Congress granted to the Executive branch almost continuously from 1974-1994 and then again as "Trade Promotion Authority" from 2002-2007.

[8] See KRW Section 5.1 for more details.

Rather than developing a detailed model of the DSP, KRW treat the DSP as a "black box," where the key feature is that settlement occurs with delay. For a period that begins in the D state, the dispute is resolved, and the state is switched to C, with probability p where p is exogenous and is meant to capture the idea that dispute resolution may entail costs including delay. I will follow an alternative and equivalent convention by assuming that the state is switched back to the "cooperative state" T periods after a dispute is initiated.

Thus the timing of actions is the following. If the countries are in state C at the start of period t, they choose any agreement that is supportable in state C and communicate the agreement to the DSI. As long as their tariff choices adhere to the agreement, they remain in state C at the start of period $t + 1$. If one or both countries defect from the agreement, however, a dispute arises, and the state is switched to D at the start of period $t + 1$.

If the countries are in state D at the start of period t, the state will only be switched back to C if $t - 1$ was the Tth period since the beginning of the dispute. In this case, the countries immediately negotiate an agreement supportable in the cooperative state and communicate it to the DSI. If the dispute is unresolved, the state remains D through the start of period $t + 1$, irrespective of what tariffs the countries select in the current period. In this event, the countries choose an agreement from among those that are supportable in the dispute state.

KRW define a recurrent agreement to be a subgame perfect equilibrium in which, in each period, the continuation value is consistent with this theory of negotiation. This requires, first, that countries agree to do as well as possible in each state; and second, that agreement is recurrent, in that continuation payoffs are always drawn from those that are supportable in the current state, but the countries are unable to alter the state as part of their agreement. The solution concept I employ here is that of the maximal recurrent agreement; that is, the recurrent agreement that maximizes the welfare of the executives, who I assume for simplicity are social welfare maximizers.

3.2.2 Trade Agreements with External Enforcement

Standard repeated-game models have one player in each country; here there are three, each with distinctive roles that mirror those laid out in the stage game. To review, in each period, the executives can re-negotiate the trade agreement, the legislature can break the trade agreement and set trade-war tariffs, and the lobby can choose whether or not to exert lobbying effort and how much effort to exert.

Given that each trading partner submits to the DSI, we can determine the repeated game incentives in each state. The executives will jointly maximize social welfare given the state, but have no opportunity to affect the state other than to choose tariffs that are supportable. Thus the executives maximize joint welfare subject to the incentive constraints of the other players. Again, because of symmetry and separability, it suffices to restrict attention to the home country.

In state D, no action of either the lobby or the legislature will change the state; that is, the continuation payoffs will be the same regardless of their actions. When x periods of punishment are remaining, the supportability condition for the legislature to adhere to any tariff $\boldsymbol{\tau^D}$, given the level of lobbying effort e, is

$$W(\gamma(e), \boldsymbol{\tau^D}) + \delta V_{ML}^D \geq W(\gamma(e), \tau^R(e, \tau^{*D})), \tau^{*D}) + \delta V_{ML}^D$$

where V_{ML}^D is the continuation value of the median legislator in the dispute state and $\tau^R(\tau^{*D}))$ is the home legislature's best response to τ^{*D} given $\gamma(e)$.[9] Since future payoffs will not be impacted by current actions, the legislature has no incentive to choose anything other than its static best response. Thus the only tariffs that can be supported in state D are the static best responses to the lobby's choice e.

The lobby faces an analogous problem. Because nothing the lobby does can impact the disposition of the DSP, it will choose the effort level that maximizes static profits. Thus the unique tariffs that are supported in state D will be identical to those in the trade-war phase of the one-shot game. I label these $\boldsymbol{\tau^{tw}} = (\tau^{tw}, \tau^{*tw})$.[10]

The supportability conditions in state C are quite different. I will assume that one legislature is randomly assigned the opportunity to break the agreement in any given

[9]Note that separability of the economy implies that the best response tariff is independent of the trading-partner's choice. In what follows I will therefore drop this dependence to simplify notation.

[10]For details, see Buzard 2012 Section 3.1.

period. This implies the following constraint on the trade agreement tariffs $\boldsymbol{\tau^a}$:

$$W(\gamma(e), \boldsymbol{\tau^a}) + \delta V_{ML}^C \geq W(\gamma(e), \tau^R(e)), \tau^{*a}) + \delta V_{ML}^D$$

where V_{ML}^C is the continuation value of the median legislator in the cooperative state. If the punishment is T periods in the dispute state (where only trade-war tariffs can be chosen), then the only part of the continuation values that need be considered are the next T periods because after T periods, the relationship will revert back to cooperation in either state and so the continuation value will be the same from period $T+1$ on. When in state C in the future, the executives will choose the same trade-agreement tariffs because they will maximize welfare subject to the same supportability conditions; in state D, the argument above shows that $\boldsymbol{\tau^{tw}}$ must be chosen. Therefore we have[11]

$$W(\gamma(e), \boldsymbol{\tau^a}) + \frac{\delta - \delta^{T+1}}{1-\delta} W(\gamma(e), \boldsymbol{\tau^a}) \geq W(\gamma(e), \tau^R(e)), \tau^{*a}) + \frac{\delta - \delta^{T+1}}{1-\delta} W(\gamma(e), \boldsymbol{\tau^{tw}}) \tag{3.4}$$

and

$$\pi(\tau^a) + \frac{\delta - \delta^{T+1}}{1-\delta} \pi(\tau^a) \geq \pi(\tau^R(e)) + \frac{\delta - \delta^{T+1}}{1-\delta} \left[\pi(\tau^{tw}) - e_{tw}\right] - e_b \tag{3.5}$$

Because the countries will not be able to do anything to change the disposition of the DSI after a dispute has been triggered, it is only these constraints for T-length punishments that must be checked; once a punishment has been triggered, the dispute-state incentive conditions are the relevant ones.

3.3 Trade Agreement Structure

We can write the executives' joint problem as

$$\max_{\boldsymbol{\tau^a}} \frac{\boldsymbol{W_E(\tau^a)}}{1-\delta} \quad \text{subject to } (3.7) \text{ and } (3.8) \tag{3.6}$$

$$\frac{\delta - \delta^{T+1}}{1-\delta} \left[W(\gamma(e), \boldsymbol{\tau^a}) - W(\gamma(e), \boldsymbol{\tau^{tw}})\right] \geq W(\gamma(e), \tau^R(e)), \tau^{*a}) - W(\gamma(e), \boldsymbol{\tau^a}) \tag{3.7}$$

$$e_b \geq \pi(\tau^R(e_b)) - \pi(\tau^a) + \frac{\delta - \delta^{T+1}}{1-\delta} \left[\pi(\tau^{tw}) - e_{tw} - \pi(\tau^a)\right] \tag{3.8}$$

Where Inequalities 3.7 and 3.8 are simple rearrangements of 3.4 and 3.5.

[11] Note that $\delta + \delta^2 + \ldots + \delta^i = \sum_{t=1}^{i} = \sum_{t=1}^{\infty} - \sum_{t=i+1}^{\infty} = \frac{\delta}{1-\delta} - \frac{\delta^{i+1}}{1-\delta} = \frac{\delta - \delta^{i+1}}{1-\delta}$.

To understand how the executives optimally structure trade agreements, we must first examine the incentives of the lobbies and how the legislatures make decisions regarding breach of the trade agreement. The symmetric structure of the model permits restriction of attention to the home country.

I will consider the economically interesting case in which, for a given δ and T, there exists a non-trivial trade agreement in the absence of lobbying, that is, one in which the lowest supportable cooperative tariffs are strictly lower than the trade-war (i.e. non-cooperative) level. Call the trade-agreement tariffs in the absence of lobbies $\boldsymbol{\tau}^{\boldsymbol{a}}_{NL}$. If $\boldsymbol{\tau}^{\boldsymbol{a}}_{NL} = \boldsymbol{\tau}^{\boldsymbol{tw}}$, the lobby has no incentive to be active and the extra constraint implied by the presence of the lobby does not bind.

In state C, the lobby has a two-stage problem. First, for the given $\boldsymbol{\tau}^{\boldsymbol{a}}$, δ and T, it calculates the minimum e_b required to induce the legislature to break the trade agreement. Call this minimum effort level $\overline{e}(\boldsymbol{\tau}^{\boldsymbol{a}})$. This calculation of precise indifference is possible because I have assumed here that the political process is certain—that is, all actors know precisely how lobbying effort affects the identity of the median legislator through $\gamma(e)$.

The e_b required to break the agreement will produce a "cheater's payoff" of $\pi(\tau^R(\overline{e}))$. The lobby will then compare its current and future payoffs from inducing a dispute net of lobbying effort (that is, $\pi(\tau^R(\overline{e})) + \delta V_L^D - \overline{e}$) to the profit stream from the trade agreement with no lobbying effort ($\pi(\tau^a) + \delta V_L^C$). With the appropriate substitutions and rearrangements, this is just Condition (3.8) evaluated at $\overline{e}$. If the former is larger, it induces the cheapest possible break; if the latter is larger, the lobby chooses to be inactive and the agreement remains in force.

The executives maximize social welfare by choosing the lowest tariffs such that the trade agreement they negotiate remains in force. Thus they must raise tariffs to the point that makes the lobby indifferent between exerting effort $\overline{e}(\boldsymbol{\tau}^{\boldsymbol{a}})$ and disengaging completely,[12] provided that this also satisfies the legislative constraint. By construction, the legislative constraint will always be satisfied. Because $\overline{e}(\boldsymbol{\tau}^{\boldsymbol{a}})$ is calculated to make the median legislator indifferent between cooperating and initiating a dispute, when the lobby is disengaged ($e_b = 0$) the median legislator cannot prefer to break the agreement

[12] Here I assume that the lobby does not exert effort when indifferent; if one were to assume the opposite, tariffs would have to be raised an extra ε.

since her preferred tariff is lower than at $\overline{e}(\boldsymbol{\tau^a})$.

Lemma 9 in the Appendix demonstrates that the solution to the executives' problem is well-defined. This combined with the immediately preceding discussion demonstrates the following result.

Result 5. *In the case of political certainty, the equilibrium trade agreement will induce zero lobbying effort and will never be subject to dispute. The executives will choose the minimum tariff level that induces the lobby to choose* $e_b = 0$*.*

At the equilibrium tariffs, the lobby's constraint will bind, while the legislature's will not. The cost of provoking a dispute, however, is derived from the legislature's constraint, which is then made slack when the lobby is disengaged.

3.4 Optimal Dispute Resolution

In an environment without lobbying, KRW show that social welfare increases (that is, trade-agreement tariffs can be reduced) as punishments are made stronger. This can be seen here if we restrict attention to the legislature's constraint:

$$\frac{\delta - \delta^{T+1}}{1-\delta}\left[W(\gamma(e), \boldsymbol{\tau^a}) - W(\gamma(e), \boldsymbol{\tau^{tw}})\right] \geq W(\gamma(e), \tau^R(e)), \tau^{*a}) - W(\gamma(e), \boldsymbol{\tau^a})$$

This constraint is made less binding as T increases—that is, as we raise the number of periods of punishment. The intuition is straightforward: the per-period punishment is felt for more periods as the one period of gain from defecting remains the same. Thus larger deviation payoffs can be supported as T increases.

Lemma 7. *The slackness of the legislative constraint is increasing in* T*.*

Thus the environment with no lobby gave no model-based prediction about the optimal length of punishment. Longer is better, but there are renegotiation constraints that must be taken into account that are outside of the model as well as other concerns.

The lobby's constraint works in the opposite direction in relation to T:

$$e_b \geq \pi(\tau^R(e_b)) - \pi(\tau^a) + \frac{\delta - \delta^{T+1}}{1-\delta}\left[\pi(\tau^{tw}) - e_{tw} - \pi(\tau^a)\right]$$

Here, the lobby benefits in each dispute period, and so the total profit from a dispute is increasing in T. Thus we have

Lemma 8. *The slackness of the lobbying constraint is decreasing in T.*

The interaction of the impact of the length of the punishment on these two constraints is quite nuanced; in many cases, adding the lobbying constraint provides a prediction for the optimal T.

As the executives choose the smallest $\boldsymbol{\tau^a}$ that makes the lobby indifferent at $\overline{e}(\boldsymbol{\tau^a})$, we must analyze the lobby's constraint (Expression 3.8) evaluated at $\overline{e}(\boldsymbol{\tau^a})$ to determine the optimal T. The derivative of this constraint with respect to T is

$$\frac{-\frac{\delta^{T+1} ln\delta}{1-\delta}\left[W(\gamma(\overline{e}),\boldsymbol{\tau^a})-W(\gamma(\overline{e}),\boldsymbol{\tau^{tw}})\right]}{\frac{\partial\gamma}{\partial e}\left[\pi(\tau^R(\gamma(\overline{e})))-\pi(\tau^a)\right]+\frac{\delta-\delta^{T+1}}{1-\delta}\frac{\partial\gamma}{\partial e}\left[\pi(\tau^{tw})-\pi(\tau^a)\right]}+\frac{\delta^{T+1}ln\delta}{1-\delta}\left[\pi(\tau^{tw})-e_{tw}-\pi(\tau^a)\right] \qquad (3.9)$$

If this expression is negative for all T, the constraint is most slack at $T = 0$, which might seem to be at odds with incentivizing cooperation by the legislature; however, unless the legislature is so biased toward the lobbying industry that its preferred tariff at $e = 0$ is above the trade-agreement level (i.e. the trade agreement level that is required to disengage the lobby), the legislature will have no incentive to defect from the agreement.

On the other hand, if this expression is positive for all T, the constraint is most slack as T approaches infinity and so we are in a case similar to that of the model without lobbying where a ad-hoc renegotiation constraint determines the upper bound on the punishment length. Here, the legislative constraint outweighs concerns about provoking lobbying effort. Perhaps most interesting are intermediate cases where the optimal T is interior—that is, the punishment length optimally balances the need to punish legislators for deviating with that of not rewarding lobbies too much for provoking a dispute.

The intuition is clearest if we examine the case of perfectly patient actors, that is, let $\delta \to 1$. This essentially removes the influence of the period of cheater's payoffs in which the interests of the legislature and the lobby are aligned (both do better in the defection stage) and exposes the differences between them in the dispute phase. In the limit, the derivative of the constraint with respect to T becomes

$$\frac{W(\gamma(\overline{e}),\boldsymbol{\tau^a})-W(\gamma(\overline{e}),\boldsymbol{\tau^{tw}})}{\frac{\partial\gamma}{\partial e}\left\{\left[\pi(\tau^R(\gamma(\overline{e})))-\pi(\tau^a)\right]+T\left[\pi(\tau^{tw})-\pi(\tau^a)\right]\right\}}-\left[\pi(\tau^{tw})-e_{tw}-\pi(\tau^a)\right]$$

$\overline{e}$ is determined so that $W(\gamma(\overline{e}),\boldsymbol{\tau^a})-W(\gamma(\overline{e}),\boldsymbol{\tau^{tw}})$ is always positive, so the numerator of the fraction is positive. The trade-agreement tariff is always lower than both the trade war tariff and the best-response (cheater's) tariff and $\frac{\partial\gamma}{\partial e}$ is positive by Assumption

3, so the denominator is always positive. Note that the only influence of T on the entire expression is through this denominator, so the value of the expression is decreasing in T.

The second term, the change in the lobby's gain from a break in the trade agreement, can for extremely large values of $\boldsymbol{\tau}^{\boldsymbol{a}}$ be negative. If this is so, the entire expression is positive and the optimal T is the largest value possible. Intuitively, the lobby has nothing to gain from causing a break, so the legislature's incentives are the only ones of concern.

In the case of interest where the lobby potentially has an interest in breaking the agreement, the right-hand term is positive. Here where we've taken $\delta \to 1$, the rate of change of the lobby's gain is constant.

Depending on the relative magnitudes, the overall expression may be positive for small T and then become negative, or it may be negative throughout. In the former case, the optimal T will be interior, while in the latter it will be zero. The expression cannot be positive for all values of T, so it cannot be optimal to have arbitrarily long punishments when the players approach perfect patience.

Result 6. *In the case of political certainty with perfectly patient players, the optimal punishment scheme precisely balances the future incentives of the lobby and legislature. It always lasts a finite number of periods and may be of zero length if the influence of lobbying on legislative preferences is extraordinarily strong ($\frac{\partial \gamma}{\partial e}$ is sufficiently high).*

The key intuition for distinguishing between the situations described in Result 6 comes from examining the properties of the political process. If $\frac{\partial \gamma}{\partial e}$ is moderate, the positive term is more likely to dominate in the beginning and lead to an interior value for the optimal T, whereas extremely large values for $\frac{\partial \gamma}{\partial e}$ make it more likely that the boundary case occurs. For a given effort level, this derivative will be smaller when the lobby is less influential; that is, when a marginal increase in e creates a smaller increase in the legislature's preferences. Thus when the lobby is less powerful ($\frac{\partial \gamma}{\partial e}$ is smaller), longer punishments are desirable. If the lobby is very influential, the same length of punishment will have a larger impact on the legislature's decisions (the impact on the gain accruing to the lobby does not change). This tips the balance in favor of shorter punishments.

This intuition generalizes for all δ as in Expression (3.9). Here the second-order condition is more complicated and can be positive if $\frac{\partial \gamma}{\partial e}$ is very small. That is, if the lobby has very little influence in the legislature, it is conceivable that welfare will be maximized by making T arbitrarily large.

Result 7. *In the case of political certainty, if non-trivial cooperation is possible in the presence of a lobby, the optimal punishment scheme is finite when the influence of lobbying on legislative preferences is sufficiently strong ($\frac{\partial \gamma}{\partial e}$ is sufficiently high).*

This helps to complete the comparison to the standard repeated-game model without lobbying. There, grim-trigger (i.e. infinite-period) punishments are most helpful for enforcing cooperation (cfr. KRW's Proposition 4). I have shown here that shorter punishments are most often optimal in the presence of lobbying. This is because long punishments incentivize the lobby to exert more effort to break trade agreements.

However, the model with no lobbies and one with very strong lobbies can be seen as two ends of a spectrum parameterized by the strength of the lobby. The optimal punishment will lengthen as the political influence of the lobby wanes and the desire to discipline the legislature becomes more important relative to the need to de-motivate the lobby.

3.5 Conclusion

I have integrated a separation-of-powers policy-making structure with lobbying into a theory of recurrent trade agreements. This theory takes seriously the idea that the threat of renegotiation can undermine punishment when cooperation is meant to be enforced through repeated interaction alone. Assuming that countries can bind themselves to condition their negotiations on the state designation of a dispute settlement institution allows punishments to become incentive compatible.

I have shown that, given complete information about the outcome of the lobbying and political process, the executives maximize social welfare by choosing the lowest tariffs that make it unattractive for the lobbies to exert effort toward provoking a trade dispute. Thus the problem with the lobby adds this extra constraint to the standard problem. While the constraint on the key repeated-game player, which here is the legislature, is loosened by increasing the punishment length, this new constraint due to the presence of lobbying becomes tighter as the punishment becomes more severe. This happens because the lobby *prefers* punishment periods in which tariffs, and thus its profits, are higher. It thus has increased incentive to exert effort as the punishment lengthens.

In a model with only the legislature, welfare increases with the punishment length. Here, this result only occurs if the lobby is sufficiently weak. As the lobby's political influence grows, the optimal punishment length becomes shorter: in the race between incentivizing the legislature and the lobby, the need to de-motivate the lobby begins to win. This suggests that a key consideration when designing the length of dispute settlement procedures is how to optimally balance the incentives of those capable of breaking trade agreements with the political forces who influence them, *given* the strength of that influence.

Key next steps include exploring the welfare implications of alternative dispute settlement procedures as well as the impact of political uncertainty on their optimal design.

3.6 Appendix

Lemma 9. *A solution to the executives' problem (3.6) exists for all δ and all T.*

Proof: The executives' problem is to minimize $\boldsymbol{\tau^a}$ such that both the legislature's and the lobby's constraints are satisfied. If the solution to the problem in the absence of lobbies (i.e. with only the legislature's constraint, and that evaluated at $e_b = 0$) cannot be satisfied for any $\boldsymbol{\tau^a} < \boldsymbol{\tau^{tw}}$ (that is, $\boldsymbol{\tau^a_{NL}} = \boldsymbol{\tau^{tw}}$), then the solution to (3.6) will also be $\boldsymbol{\tau^{tw}}$.

Consider the case where $\boldsymbol{\tau^a_{NL}} < \boldsymbol{\tau^{tw}}$. I rewrite the constraints with the payoffs normalized and $\delta = \mathrm{e}^{-r\Delta}$ where r is the interest rate and Δ is the period length:

$$\left(1 - \mathrm{e}^{-r\Delta}\right)\left[W(\gamma(e), \tau^R(e)), \tau^{*a}) - W(\gamma(e), \boldsymbol{\tau^a})\right] - \\ \mathrm{e}^{-r\Delta}\left(1 - \mathrm{e}^{-r\Delta T}\right)\left[W(\gamma(e), \boldsymbol{\tau^a}) - W(\gamma(e), \boldsymbol{\tau^{tw}})\right] \geq 0 \quad (3.10)$$

$$\left(1-\mathrm{e}^{-r\Delta}\right)e_b \\ -\left(1-\mathrm{e}^{-r\Delta}\right)\left[\pi(\tau^R(e_b))-\pi(\tau^a)\right]-\mathrm{e}^{-r\Delta}\left(1-\mathrm{e}^{-r\Delta T}\right)\left[\pi(\tau^{tw})-e_{tw}-\pi(\tau^a)\right]\geq 0 \tag{3.11}$$

The proof is via the Intermediate Value Theorem. I take as the leftmost boundary $\boldsymbol{\tau}^{\boldsymbol{a}}_{NL}$ (the trade agreement chosen by the executives when there is no lobby) because this is the lowest possible tariff the executives can achieve before the additional constraint implied by the presence of lobbies is added. By construction, $\overline{e}(\boldsymbol{\tau}^{\boldsymbol{a}}_{NL})=0$. The gain to the lobby of a break in the trade agreement here is

$$\left(1-\mathrm{e}^{-r\Delta}\right)\left[\pi(\tau^R(0))-\pi(\tau^a_{NL})\right]+\mathrm{e}^{-r\Delta}\left(1-\mathrm{e}^{-r\Delta T}\right)\left[\pi(\tau^{tw})-e_{tw}-\pi(\tau^a_{NL})\right]$$

If this gain is non-positive, the lobby has no incentive to exert effort and so $\boldsymbol{\tau}^{\boldsymbol{a}}_{NL}$ is the solution to the executives' problem.[13] If the lobby's gain is strictly positive, the left-hand size of Expression 3.11 is negative. The executives will have to raise the trade-agreement tariff to prevent the agreement from being broken in this case because the lobby's constraint is not satisfied.

Next, look at the rightmost boundary, that is $\boldsymbol{\tau^a}=\boldsymbol{\tau^{tw}}$. The executives have no incentive to set a trade-agreement tariff above the trade-war (i.e. non-cooperative) level, as this is the highest tariff-level they will have to face even if the trade-agreement is broken. When $\boldsymbol{\tau^a}=\boldsymbol{\tau^{tw}}$, the legislature's constraint becomes

$$\left[W(\gamma(e_b),\tau^R(e_b),\tau^{*tw})-W(\gamma(e_b),\boldsymbol{\tau^{tw}})\right]\geq 0$$

To make this condition hold with equality, we need e_b so that $\tau^R(e_b)=\tau^{tw}$; that is, $\overline{e}=e_{tw}$.

The lobby's gain will then be

$$\left(1-\mathrm{e}^{-r\Delta}\right)\left[\pi(\tau^{tw})-\pi(\tau^{tw})\right]+\mathrm{e}^{-r\Delta}\left(1-\mathrm{e}^{-r\Delta T}\right)\left[\pi(\tau^{tw})-e_{tw}-\pi(\tau^{tw})\right]= \\ -\mathrm{e}^{-r\Delta}\left(1-\mathrm{e}^{-r\Delta T}\right)e_{tw} \tag{3.12}$$

[13] Note, in particular, that this is the case if both Δ and $\Delta T\to 0$.

Therefore, at $\boldsymbol{\tau^a} = \boldsymbol{\tau^{tw}}$, the left-hand side of Expression (3.11) is

$$\left(1 - \mathrm{e}^{-r\Delta}\right) e_{tw} + \mathrm{e}^{-r\Delta} \left(1 - \mathrm{e}^{-r\Delta T}\right) e_{tw}.$$

This is always positive, which only requires $\Delta T > 0$ (recall that the special case of both Δ and $\Delta T \to 0$ has already been treated above).

In order to apply the Intermediate Value Theorem, it is left to show that the left-hand side of Expression (3.11) is continuous in $\boldsymbol{\tau^a}$. The lobby's gain is continuous in the tariffs by the assumptions on profits in Section 3.1. Given Assumption 3 and the assumptions in Section 3.1, $\overline{e}$ is continuous by the Implicit Function Theorem. Because the sum of continuous functions is continuous, we have the desired result and the Intermediate Value Theorem can be applied to ensure that, under the conditions stated above, the left-hand side of Expression (3.11) attains zero on the interval $\left[\boldsymbol{\tau}^{\boldsymbol{a}}_{NL}, \boldsymbol{\tau^{tw}}\right]$ at least once; the solution to Problem 3.6 is at the minimum such $\boldsymbol{\tau}$. ∎

3.7 References

Bagwell, Kyle and Robert Staiger (1990), "A Theory of Managed Trade," *American Economic Review*, 80, 779-795.

Bagwell, Kyle and Robert Staiger (1997a), "Multilateral Tariff Cooperation During the Formation of Customs Unions," *Journal of International Economics*, 42, 91-123.

Bagwell, Kyle and Robert Staiger (1997b), "Multilateral Tariff Cooperation During the Formation of Free Trade Areas," *International Economic Review*, 38, 291-319.

Bagwell, Kyle and Robert Staiger (2002), "The Economics of World Trading System," MIT Press, Cambridge.

Bagwell, Kyle and Robert Staiger (2005), "Enforcement, Private Political Pressure, and the General Agreement on Tariffs and Trade/World Trade Organization Escape Clause," *Journal of Legal Studies*, 34, 471-513.

Beshkar, Mostafa (2011), "Arbitration and Renegotiation in Trade Agreements," Available at http://dl.dropbox.com/u/8414554/TariffBargaining.pdf.

Buzard, Krisy (2012), "Trade Agreements, Lobbying and Separation of Powers," available at: http://dss.ucsd.edu/~kbuzard/pdfs/Separation_of_Powers.pdf.

Cotter, Kevin. D. and Shannon K. Mitchell (1997), "Renegotiation-Proof Tariff Agreements," *Review of International Economics*, 5, 348-371.

Destler, I.M. (2005), "American Trade Politics," Institute for International Economics, 4th ed.

Dixit, Avinash (1987), "Strategic Aspects of Trade Policy," In: Bewley, Truman F. (Ed.), Advances in Economic Theory: Fifth World Congress. Cambridge University Press, Cambridge, pp. 329-362.

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

Grossman, Gene and Elhanan Helpman (1994), "Protection for Sale," *The American Economic Review*, 84, 833-850.

Grossman, Gene and Elhanan Helpman (1995), "Trade Wars and Trade Talks," *The Journal of Political Economy*, 103, 675-708.

Klimenko, Mikhail, Garey Ramey and Joel Watson (2008), "Recurrent Trade Agreements and the Value of External Enforcement," *Journal of International Economics*, 74, 475-499.

Kovenock, Dan and Marie Thursby, (1992), "GATT, Dispute Settlement and Cooperation," *Economics and Politics*, 4, 151-170.

Lohmann, Susanne and Sharyn OHalloran (1994), "Divided Government and US Trade Policy," *International Organization*, 48, 595-632.

Ludema, Rodney D. (2001), "Optimal International Trade Agreements and Dispute Settlement Procedures," *European Journal of Political Economy*, 17, 355-376.

Maggi, Giovanni (1999), "The Role of Multilateral Institutions in International Trade Cooperation," *American Economic Review*, 89, 190-214.

Maggi, Giovanni and Robert Staiger (2011), "Optimal Design of Trade Agreements in the Presence of Renegotiation," Available at http://www.stanford.edu/~rstaiger/Breach&Remedies_090611.pdf.

McMillan, John (1986), "Game Theory in International Economics," Harwood, New York.

McMillan, John (1989), "A Game-theoretic View of International Trade Negotiations: Implications for the Developing Countries," In: Whalley, John (Ed.), Developing Countries and the Global Trading System, vol. 1. University of Michigan Press, Ann Arbor, MI, pp. 26-44.

Milner, Helen and B. Peter Rosendorff (1997), "Democratic Politics and International trade Negotiations: Elections and Divided Government as Constraints on Trade Liberalization," *Journal of Conflict Resolution*, 41, 117-147.

Riezman, Raymond (1991), "Dynamic Tariffs with Asymmetric Information," *Journal of International Economics*, 30, 267-283.

Song, Yeongkwan (2008), "Protection for Sale: Agenda-Setting and Ratification in the Presence of Lobbying," *Korea Institute for International Trade Policy Working Paper Series* Vol. 2008-22, August 2008.