

UCLA

UCLA Electronic Theses and Dissertations

Title

Why Comply? Domestic Politics and the Effectiveness of International Courts

Permalink

<https://escholarship.org/uc/item/42k8p6xw>

Author

Peritz, Lauren J.

Publication Date

2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Why Comply? Domestic Politics and the Effectiveness of International Courts

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

by

Lauren J. Peritz

2015

© Copyright by
Lauren J. Peritz
2015

ABSTRACT OF THE DISSERTATION

Why Comply? Domestic Politics and the Effectiveness of International Courts

by

Lauren J. Peritz

Doctor of Philosophy in Political Science

University of California, Los Angeles, 2015

Professor Leslie Nicole Johns, Co-chair

Professor Arthur A. Stein, Co-chair

This dissertation asks: when do international courts promote cooperation among countries? I argue that international courts can successfully restore economic relations between disputing governments but their impact depends on domestic politics. When confronted with an adverse legal ruling from an international court, a defendant government must determine whether and when to comply. Governments are constrained by domestic institutional divisions and partisan conflict: “veto points.” Countries with substantial divisions are less likely to comply because more political actors must coordinate to implement the ruling. As partisan divisions grow, government leaders are constrained by their domestic opposition and compliance becomes more difficult. The design of the international court contributes to this effect. Courts vary in their ability to sanction violations. When the court is designed to be flexible, imposing low costs for noncompliance, the impact of domestic politics is particularly pronounced.

These arguments are tested with international trade disputes at the World Trade Organization (WTO) and the European Court of Justice (ECJ). The first empirical chapter uses WTO disputes to examine the impact of domestic politics in the defendant country on compliance with adverse legal rulings. Adverse rulings require

a defendant government to remove trade barriers so this chapter assesses compliance using trade flows. I build a novel data set on compliance using the method of synthetic case control and product-level time-series trade data. I infer the defendant complied if trade flows increased after the dispute, relative to estimated levels that would have occurred in the absence of the ruling. The estimates show compliance problems are both widespread and systematically linked to domestic politics. Domestic constraints—measured in terms of veto points—hinder compliance.

The second empirical chapter tests my main argument on the European Court of Justice. I show that domestic political constraints in European Union countries also impact compliance with adverse legal rulings. I focus on infringement disputes over trade-related issues, instances in which European member states imposed illegal barriers to intra-European commerce. This chapter uses a hierarchical model that captures the multi-level structure of the data. By examining intra-European trade over time, I show that adverse rulings lead to a modest increase in trade but this tendency is conditional on domestic politics. Defendant governments with many veto players appear impervious to adverse rulings. The findings indicate that ECJ rulings can prompt governments to open their markets to more European commerce, but that domestic politics can obstruct this process.

The third empirical chapter evaluates the effectiveness of international dispute settlement along a different dimension: the time to resolve a dispute. Because prolonged lawsuits can buy defendants time to “cheat” at the expense of plaintiffs and other members of the international institution, they can have deleterious effects on cooperation that are similar to noncompliance. This chapter demonstrates that WTO and ECJ lawsuits against defendants with many domestic veto points lasted longer on average, before the countries acquiesced. Moreover, the ill effect of veto players on dispute resolution has been stronger in the WTO than the ECJ. I argue that the design of the international court mediates the impact of domestic veto players on dispute duration.

In sum, my dissertation shows that international courts can successfully promote economic cooperation between countries but their effectiveness hinges on domestic politics.

The dissertation of Lauren J. Peritz is approved.

Richard Steinberg

Arthur A. Stein, Committee Co-chair

Leslie Nicole Johns, Committee Co-chair

University of California, Los Angeles

2015

To Judy, Leigh, and David, for your unconditional support.

TABLE OF CONTENTS

1	Introduction	1
1.1	Motivation	1
1.2	Domestic Politics of International Cooperation	3
1.3	Compliance and Effectiveness	5
1.4	Research Design, Case Selection, and Estimation	9
1.5	Overview of the Dissertation	13
2	When are International Institutions Effective? The Impact of Domestic Veto Players on Compliance with WTO Rulings	16
2.1	Introduction	16
2.2	Cooperation, Compliance, and the WTO	20
2.3	A Theory of Domestic Veto Players and Compliance	24
2.3.1	Domestic Preferences, Power, and Institutions	24
2.3.2	International Cooperation and Compliance	26
2.3.3	Illustrations	29
2.4	Assessing Compliance with WTO Rulings	32
2.4.1	Measurement Strategy	32
2.4.2	Product-Level Trade Flows	35
2.4.3	Synthetic Control Method to Measure Compliance	38
2.4.4	Methodological Advantages	43
2.5	Analysis and Results	44
2.5.1	Dependent Variable - Compliance	44
2.5.2	Explanatory Variables and Controls	46

2.5.3	Results: Domestic Veto Points Hinder Compliance	47
2.5.4	Robustness	54
2.6	Conclusion	58
3	The European Court of Justice and EU Economic Cooperation: Do Domestic Political Divisions Hinder Reform?	61
3.1	Introduction	61
3.2	Regulation of Trade, ECJ Rulings, and Domestic Obstacles	63
3.2.1	Regulation of Trade	63
3.2.2	ECJ Infringement Disputes	65
3.2.3	Domestic Constraints	69
3.3	Data and Model	71
3.3.1	Dependent Variable: Intra-EU Trade	71
3.3.2	Explanatory Variables: ECJ Rulings and Domestic Veto Players	73
3.3.3	Control Variables	76
3.3.4	Model Estimation	77
3.4	Results - Classical Approach	78
3.5	Results - Bayesian Approach	83
3.5.1	Random Intercepts Model	83
3.5.2	Random Slopes Model	86
3.5.3	Robustness	88
3.6	Discussion	90
4	Comparing International Courts: How Institutional Design Mediates the Domestic Politics of Dispute Settlement	93

4.1	Introduction	93
4.2	Comparing International Courts	97
4.2.1	International Courts	97
4.2.2	The WTO Dispute Settlement Mechanism	99
4.2.3	The European Court of Justice	102
4.2.4	Comparing Institutional Design	104
4.3	Institutional Design and the Domestic Politics of International Coop- eration	108
4.3.1	Domestic Politics of International Cooperation	108
4.3.2	Relationship Between Domestic Politics and Institutional Design	110
4.4	Empirical Analysis	112
4.4.1	Data Sources	112
4.4.2	Disputes Stages and Veto Players	113
4.4.3	Dispute Duration and Veto Players	116
4.4.4	Illustrations	123
4.5	Discussion	127
5	Conclusion	129
5.1	Domestic Politics and the Effectiveness of International Courts	129
5.2	Summary	133
5.3	Empirical Contributions and Theoretical Significance	134
5.4	Questions and Implications	138
A	Technical Material for Analysis of WTO Disputes	145

A.1	Measuring Compliance with the Synthetic Control Method	145
A.2	Trade Flows Used in Synthetic Control Method	150
A.3	Robustness Checks for Selection Effects	151
B	Technical Material for Analysis of ECJ Disputes	155
B.1	Code for Bayesian Analysis	155
	Bibliography	157

LIST OF FIGURES

2.1	Hypothetical Compliance and Noncompliance Trade Patterns	34
2.2	WTO Disputes with Compliance and Noncompliance Trade Patterns	42
2.3	Predicted Effect of Veto Points on Compliance	51
2.4	Robustness to Dispute Timing	57
3.1	Intra-EU Imports Over Time by Country	73
3.2	Parameter Estimates for Country-Random Intercept Models	84
3.3	Country-Random Intercepts	85
3.4	Parameter Estimates Varying Temporal Lag	86
3.5	Country-Random Slopes on Interaction Term and Shared Parameter Estimates for Other Covariates	87
3.6	Coefficient Estimates for Placebo Test	89
4.1	Predicted Survival of Legal Disputes with Domestic Veto Points	122
A.1	Schematic of Trade Flows	151

LIST OF TABLES

2.1	Probit Models for Compliance with WTO Rulings, 1995–2011	49
2.2	Ordered Probit Models and Linear Model of Compliance with WTO Rulings, 1995–2011	53
3.1	Frequency of Adverse ECJ Rulings per Year By Country	74
3.2	Linear Models for EU-14 Import Share with Various Lags, 1988-2005	80
3.3	Linear Models for EU-14 Import Share with Various Lags and Eco- nomic Controls and AR(1) Correction, 1988-2005	81
3.4	Linear Models for EU-14 Import Share with Political Controls, 1988- 2005	82
4.1	Disputes at WTO, 1995-2012 and ECJ Infringement Disputes, 1978- 1999, by Stage of Legal Proceeding	114
4.2	Cox Proportional Hazard Model of WTO Disputes, 1995-2012, where Event is Dispute Resolution	118
4.3	Cox Proportional Hazard Model of ECJ Disputes, 1978-1999, where the Event is Dispute Resolution	120
A.1	Two-Step Heckman Selection Models for WTO Disputes, 1995-2011 .	154

ACKNOWLEDGMENTS

This dissertation would not have been possible without the generous support and guidance I received from my advisers, Leslie Johns and Arthur Stein. From my first to last day at UCLA, Leslie has been the most dedicated and thoughtful mentor I could have imagined. Always available to discuss ideas, critique drafts, and quell my self-doubt, Leslie has pushed me to transform from an enthusiastic student into a professional scholar. Over the years, Art has dedicated countless hours to our discussions about my research and the profession. He has always encouraged me to think creatively about the “big” questions in international relations. To both, I am indebted and deeply grateful. The other members of my dissertation committee, Jeffrey Lewis, Ronald Rogowski, and Richard Steinberg, also provided invaluable feedback along the way. Jeff taught me much of what I know about statistics and has always made the time to work with me through analytical puzzles. And Jeff also knows quite a bit about international trade disputes. Ron has helped me connect my ideas about international organizations to domestic institutions and has shared many words of wisdom about European politics. Richard has encouraged me to link the arguments in this dissertation to broader questions in international law. I am fortunate to have had the opportunity to work with each of these exceptional professors.

Numerous other faculty at UCLA have enriched my graduate education and weighed in on different aspects of this project. In particular, I thank Kathy Bawn, Jim DeNardo, Barbara Geddes, Miriam Golden, Mark Handcock, Chad Hazlett, Barry O’Neill, Michael Ross, and Mike Thies. They have guided me toward interesting research questions while teaching me the methodological skills I need to answer them. In addition to this educational support, I received generous funding from the Department of Political Science, UCLA Graduate Division, the University of California Institute on Global Conflict and Cooperation, and the National Sci-

ence Foundation.¹ Of course, none of this would have been possible without Joseph Brown, graduate adviser, who manages to keep the Department of Political Science running and greases the wheels of university bureaucracy whenever I need the help.

This project also benefited from feedback from many scholars outside my UCLA circle. I thank Chad Bown, Clifford Carrubba, Songying Fang, Barbara Koremenos, Jeffrey Kucik, Krzysztof Pelc, and Jana von Stein for their valuable comments on various iterations of this research. By taking the time out of their busy lives to give me feedback, these scholars have helped to create an engaging intellectual community and set an example I am eager to follow.

My fellow graduate students in the Department of Political Science were a source of tremendous support. I thank Ruth Carlitz for sharing every step of graduate school with me as well as Sarah Brierley, Sebastián Garrido de Sierra, Kristen Kao, Paasha Mahdavi, Felipe Nunes, Amanda Rizkallah Chown, and Andrea Vilán. It was an honor to share research ideas, half-baked work, classes, and an office with them. Their camaraderie, warmth, and humor kept me motivated, especially during the more challenging stages of dissertation research. Last and certainly not least, I am grateful to Becca DiBennardo, Erin Flinn, Will Foster, Lauren Goldstein, Mike Kennett, and Anne White, my west-coast family. With their help, I have found a wonderful home in Los Angeles and relished my six years exploring the city and surrounding mountains in fine company.

¹IGCC Dissertation Fellowship 2014-2015 and NSF Dissertation Improvement Grant SES-1421326.

VITA

Education:

- | | |
|------|---|
| 2015 | M.S. (Statistics) University of California, Los Angeles. |
| 2006 | M.A. (Government) Cornell University. |
| 2003 | B.S. (College Program in Engineering Physics) Cornell University. |

Employment:

- | | |
|-----------|---|
| 2015– | Assistant Professor of Political Science. University of California, Davis. |
| 2013–2015 | Teaching Assistant. University of California, Los Angeles. |
| 2010–2014 | Graduate Student Researcher. University of California, Los Angeles. |
| 2011–2012 | Editorial Assistant. American Political Science Review, Los Angeles. |
| 2006-2009 | Intellectual Property Technology Specialist. Wilmer, Cutler, Pickering, Hale, and Dorr LLP, Boston. |
| 2004 | Physical Science Technician. The Harvard-Smithsonian Center for Astrophysics, Cambridge. |

PUBLICATIONS

“The Design of Trade Agreements” with Leslie Johns. 2015. *Oxford Handbook of the Politics of International Trade*, Lisa Martin, ed. Oxford University Press.

CHAPTER 1

Introduction

1.1 Motivation

International agreements shape global politics. They regulate international trade and investment, enforce human rights, establish grounds for lawful military action, and create common standards for innumerable other domains of international affairs. Many of these agreements create formal international organizations such as the European Union, International Criminal Court, North Atlantic Treaty Organization, United Nations, and World Trade Organization. Countries join international organizations for a variety of reasons. One prevailing view is that these cooperative arrangements help stave off costly conflicts that arise when each country plays by its own rules. When countries commit to common principles, policy, and practice in the form of international agreements, they are less likely to find themselves in conflict and apt to enjoy more efficient outcomes: greater international trade and investment, lower risk of war, better rights for their citizens. In this sense, international institutions are thought to provide a rational solution to deep-seated coordination and collaboration problems that plague global affairs. Yet conflicts between countries inevitably arise.

In recent decades, international courts have proliferated. These institutions apply the burgeoning body of international law to resolve conflicts between national governments or between non-governmental plaintiffs and countries. Like their domestic counterparts, international courts adjudicate legal disputes and issue verdicts—

rulings which often declare the defendant government's policies inconsistent with international obligations. But unlike domestic courts, which enjoy the backing of the state, international courts cannot be sure that aggrieved governments will implement their rulings. Enforcement is scant. For example, a defendant government may dig in its heels and resist an international court that orders it to remove an illegal trade barrier, escaping retaliation.

Without consistent international enforcement, governments which bear the responsibility to faithfully implement adverse rulings sometimes fall short of their obligations. The temptation to defy international courts may be acute when national governments face particular domestic political conditions, including pressure from industry groups that have a stake in the outcome or elections that turn on the support of key constituencies. It is in these instances that governments are likely to shirk their international commitments. International courts that handle economic disputes are particularly susceptible to these types of compliance problems. Because international economic agreements often have distributive implications at home, rulings over trade, investment and related topics are especially likely to mobilize domestic opposition.¹

This dissertation asks a question fundamental to the study of international relations: *when do international courts promote economic cooperation between countries?* It is fundamental because it probes the significance of international organizations themselves. Scholars have long debated whether international institutions actually help countries cooperate or whether they are just a thin veneer over the self-help conditions that characterize Hobbes' "state of nature." International dispute settlement offers a window into the effectiveness of these institutions. These disputes highlight the behavior of governments when they do not want to follow through on their primary international commitments.

¹For example, a free trade agreement might confer advantages to consumers while removing protection for producers.

1.2 Domestic Politics of International Cooperation

In the chapters that follow, I offer a domestic political explanation for when international courts are effective, building on a rich literature that links politics within governments to international cooperation (e.g. Milner and Rosendorff 1997; Martin 2000; Rosendorff and Milner 2001; Mansfield, Milner, and Rosendorff 2002; Goldstein and Steinberg 2008; Dai 2005; Davis 2012). I argue that international courts can successfully restore economic cooperation between disputing governments but their impact depends on domestic politics. When confronted with an adverse ruling from an international court, a defendant government must determine whether and when to comply. Defendant governments are constrained by domestic institutions and partisan conflict. These constraints form “veto points.” Governments with substantial institutional divisions are less likely to comply because more political actors must coordinate to implement the legal ruling. As partisan divisions grow, government leaders are constrained by their domestic opposition and compliance becomes more difficult. All else equal, a more constrained government will be less likely to implement international court rulings that require domestic policy change.

The design of the international court contributes to this effect. Courts vary in their ability to sanction violations. Some allow defendant governments to violate their international obligations for several years without facing punishment, while others act promptly and impose retrospective fines for the entire period of the violation. When an international court is highly flexible, imposing only small or belated penalties on defiant governments, the disputes that lead to litigation tend to be long and drawn out. Instead of complying immediately, the defendant will often delay until domestic costs subside. The theory presented in the chapters that follow predicts that domestic veto points obstruct compliance with an international court’s ruling and that this effect is most prominent when the court is designed to be flexi-

ble. Together, domestic veto points and institutional design help explain patterns of compliance and timing of dispute resolution.

This project addresses a now-prominent stream of the international relations that literature looks to domestic politics to explain state behavior in global affairs. It is widely argued that democracy promotes international economic cooperation (e.g. Martin 2000; Mansfield, Milner, and Rosendorff 2002). Institutional divisions of power and partisan opposition in government—hallmarks of a functional democracy—create multiple veto points (Tsebelis, 1995). Multiple veto points narrow the set of international agreements that can be ratified, since more domestic actors must coordinate and consent to the agreement (Milner and Rosendorff, 1997). Proposed agreements that do not receive widespread support will fail at the domestic stage and never be ratified. This makes the treaties enacted by democracies more credible commitments about governments’ future behavior than those by non-democratic regimes, which remain less constrained by domestic politics. In short, veto players can lock in cooperative policies and make it difficult for governments to renege on their international promises.

However, democracies can and do break their international commitments. When domestic pressure to violate treaty terms is acute, leaders have strong incentives to shirk their obligations. The risk, of course, is that these domestic pressures cause a government to exit the treaty. Many scholars argue that institutions that allow leaders to respond to such pressure by temporarily violating their commitments are more stable over time (e.g. Staiger and Bagwell 1999; Rosendorff and Milner 2001; Bown 2002; Rosendorff 2005). By permitting some violations, flexible institutions allow members to manage temporary economic shocks or political circumstances without abandoning the institution altogether. The prevalence of flexibility mechanisms in international institutions suggests that the risk that countries renege is indeed significant.

This dissertation offers an important caveat to regime-type arguments about international economic cooperation. Just as multiple veto players can lock in international commitments, they can also lock in violations of those commitments. Once an initial violation occurs, a government with many veto players is less likely to return to cooperation. Treaty violations often confer benefits on specific domestic actors, for example illegal agriculture subsidies that help farmers and, by extension, the legislators who represent them. Domestic actors who benefit most from the violation have the greatest incentive to exercise their veto power and obstruct policy changes. As a result, governments may sometimes find themselves in a state of domestic gridlock while policies that violate their international commitments remain fixed in place.

One key implication is that when a government has many veto points, flexibility mechanisms may be less successful in generating long-term stability. By locking violations in place, domestic veto players may transform what were intended to be temporary defections into long-standing breaches of international commitments. In this sense, flexibility mechanisms may fail to restore long-term cooperation. Treaty violations by one country, in turn, undermine a delicate collaborative equilibrium, occasionally prompting other countries to break their commitments as well. In this sense, defining features of democratic politics—veto players—can actually hinder international economic cooperation. Combined with the insights of previous literature, this dissertation provides evidence that democratic governance can cut both ways, sometimes helping and other times hindering international economic cooperation.

1.3 Compliance and Effectiveness

This dissertation draws on a rich literature that explains the formation and impact of international institutions from a rational choice perspective (Stein, 1990; Keohane, 1982; Downs, Rocke, and Barsoom, 1996; Abbott and Snidal, 1998; Koremenos, Lipson, and Snidal, 2001; Rosendorff and Milner, 2001; Gilligan, 2004). I adopt the

view that countries generally create international institutions in order to solve the coordination and collaboration problems that plague international politics. International economic affairs are especially prone to collaboration problems or dilemmas of common interest which, as Stein (1990) explains, arise “when the Pareto-optimal outcome that the actors mutually desire is not an equilibrium outcome” (38). To the student of game theory, this may be most familiar as a prisoner’s dilemma scenario: countries collectively gain from cooperative arrangements but each has a unilateral incentive to “deviate” and cheat, typically at the expense of its peers. If all countries choose not to cooperate—the equilibrium outcome—they are collectively worse off than had they all cooperated—the Pareto-optimal outcome. International trade cooperation is an archetypal example. Countries have a unilateral incentive to impose trade barriers but a mutual interest in exchanging goods and services freely. Trade agreements aim to move from a suboptimal outcome of high trade barriers to a more efficient one.

Of course countries don’t always abide by treaty terms. Sometimes noncompliance is the result of incomplete information. For example, a violation may reflect ambiguous or imprecise legal obligations (Chayes and Chayes, 1993). Governments are unaware that their policies are inconsistent with their international commitments. Other times, governments are aware of obligations but behave opportunistically (Yarbrough and Yarbrough, 1987*a,b*). Governments may choose to violate whenever the benefits of doing so outweigh the expected penalties. For example, a government might knowingly impose an illegal trade barrier if it believes it can evade punishment or the trade partners who are harmed lack economic leverage. Because international enforcement is typically informal and decentralized and violations can go unnoticed, governments may make the calculated choice to shirk their commitments.

Compliance problems have broad effects. One government’s noncompliance can impose significant costs on others. For example, if a government violates its trade

agreement by imposing an illegal tariff, that decreases its market for imported goods. Producers in the other countries may notice diminished profits and broader economic harm. This undermines the collaborative equilibrium: the other states then prefer to impose trade protection, leading to less efficient exchange of goods and services. Noncompliance by a few leads to a worse outcome for many.

Because the success of an international institution hinges on the consistent cooperation of its members, these organizations deploy many strategies to improve compliance. International regimes that aim to solve collaboration problems “must specify strict patterns of behavior and ensure that no one cheats” by clarifying what constitutes cooperation versus cheating and by assuring each country that it will be able to immediately spot others’ cheating (Stein 1990, 38-9). To induce cooperation over time, international organizations punish defections, publicize violations, or use other means.

Scholars have identified two forms of compliance in the context of international relations (Fisher, 1981; Simmons, 1998). First-order compliance is when a state abides by its substantive legal obligations. For trade agreements, a government complies in the first sense if its trade policy and practice reflect substantive legal commitments, e.g. low trade barriers. When a state fails to comply with treaty terms and a dispute arises, it is often required to correct the violation. Second-order compliance is when the state adjusts its policy and practice after an initial violation, often in response to international litigation. A government complies in the second sense if it loses a dispute and corrects the violation, e.g. removes an illegal trade barrier.

When international institutions enjoy high rates of compliance, it is tempting to infer that those institutions are effective at promoting cooperation among countries (Simmons, 2000; Simmons and Martin, 2012). But as Downs, Rocke, and Barsoom (1996) famously argued, the good news about compliance isn’t necessarily good news about cooperation. A state’s decision to join an international agreement is endoge-

nous to its expectations about future compliance with that agreement. Countries make international commitments that align with their interests and when they comply, they may simply be adopting behavior they would have chosen anyway, in the absence of the institution (Downs and Rocke, 1995; Downs, Rocke, and Barsoom, 1996; Von Stein, 2005). Consequently, the major obstacle to studying institutional effectiveness is that it is difficult to separate causal impact from the selection process. This is the conundrum that makes empirical studies of first-order compliance so difficult.

Second-order compliance, by contrast, can be more revealing. Countries often find themselves in international legal disputes after violating their primary obligations. When international courts issue adverse rulings, they confirm that the defendant government has failed to first-order comply. These are instances where governments demonstrate an explicit preference to violate their primary obligations—we know this because of the adverse ruling. When despite this preference, a government complies with an adverse ruling, we can infer the institution had some influence. While selection effects still loom large—*Which cases are prosecuted? When is a court willing to issue adverse rulings?*—a case of second-order compliance is an instance in which a government changed its behavior in response to the international court. In this respect, second-order compliance lends insight into effectiveness.

This dissertation touches on broader questions about the design of effective international institutions. States strive to design international courts in ways that promote cooperation (Rosendorff, 2005; Gilligan, Johns, and Rosendorff, 2010; Carrubba, 2005; Kucik and Reinhardt, 2008; Carrubba and Gabel, 2015; Johns, 2015). This means choosing institutional features that allow governments to juggle often-competing demands from international and domestic sources while still making international dispute settlement useful.² As Johns (2015) argues, when the domestic

²International courts promote cooperation in various ways. Courts clarify ambiguous legal obligations and render verdicts on alleged violations. In some cases, they authorize punishments for

temptation to violate international agreements is small enough, “strengthening international courts boosts compliance by increasing the cost of dispute settlement” (163). But when following the rules is very difficult, “because a government faces intense political or economic pressure to violate a treaty, compliance is not feasible. A state will break the regime’s rules and either participate in dispute settlement or leave the regime altogether” (163). In these circumstances, stronger courts can sometimes destabilize international cooperation.

These institutional design choices have implications for the domestic politics of international dispute settlement. As I argue below, some design features enable domestic politics to heavily influence the dispute process and likelihood of compliance. My findings suggest that in order to design international courts that are effective in promoting international cooperation, one must account for domestic constraints within countries. Optimal designs should not only account for a state’s collective benefits and costs of international cooperation, but also a government’s domestic constraints. These constraints pose real obstacles to compliance when international promises have been broken. In Chapter 5, I reflect on these implications for designing better international courts.

1.4 Research Design, Case Selection, and Estimation

My theory has two key explanatory variables: domestic political constraints and institutional design. Accordingly, the first aim is to evaluate the impact of domestic constraints on international dispute resolution—compliance with court rulings and the duration of disputes. This is the primary focus of the chapters that follow. The second aim is to demonstrate how institutional design can mediate these effects.

treaty violations, thereby legitimizing and coordinating enforcement (Busch and Reinhardt, 2000; Johns, 2011).

The research design reflects these aims. First, I examine the relationship between domestic political constraints and compliance at each of two prominent international courts. One is the World Trade Organization (WTO) Dispute Settlement Mechanism which serves as a center-point for global trade relations. The other is the European Court of Justice (ECJ), the branch of the European Union that adjudicates disputes between states, as well as lawsuits by the European Commission and private plaintiffs against member states. For each of these institutions, defendant countries routinely lose the lawsuits and are expected to reform their domestic policies to bring them into compliance. But in both courts, the compliance process can become politically fraught and the outcome is far from certain. To evaluate the outcomes in each institution, I measure compliance by estimating the effect of a legal ruling on trade flows. I then show that domestic political divisions can decrease the probability of compliance.

Second, I compare dispute outcomes at the WTO to those at the ECJ. There are many differences between the two institutions that make direct comparisons tenuous at best. To improve comparability between the two institutions, I restrict the analysis to ECJ disputes that relate to trade, commerce, and the European internal market. This creates a similar set of cases to those at the WTO so that I can examine compliance, subject to similar incentives and political forces. In comparing the two institutions, I highlight institutional design features that appear to mediate the effect of a country's domestic politics on its international behavior. These design features and their relevance for the domestic politics of compliance are discussed detail in Chapter 4.³

I select these two institutions for several reasons. First, both international courts are highly active and adjudicate cases from a relatively diverse set of litigants. Since

³Because of the significant disparities between the two institutions, I cannot draw firm conclusions about the impact of institutional design on compliance; this aspect of the research design is amenable only to suggestive discussion.

it was founded in 1995, the WTO Dispute Settlement Mechanism has dealt with more than 471 lawsuits involving over 40 different defendant countries. The defendants span the globe and while industrialized nations are the most active litigants, developing countries increasingly enter WTO disputes. The European Court of Justice handles an even higher volume of economic disputes. Among its many cases, the ECJ has adjudicated nearly 700 complaints that directly address European economic integration. All European Union members, with the exception of only the most recent admits, have been implicated in multiple lawsuits.

Because they are highly active, the two courts are especially suitable for studying the domestic politics of compliance. Both have sufficient observations for statistical analysis. Both are reputed to have strong compliance records. And both have significant variation in the defendant governments to gain empirical traction on the main explanatory variable in my theory: domestic veto players. By contrast, other dispute settlement mechanisms are used relatively infrequently by a small set of countries.

Second, I study these courts because their rulings have clear economic implications. In both, lawsuits arise when the defendant government is accused of violating its treaty obligations. Adverse rulings typically require a defendant government follow through on those obligations by opening its economy through trade liberalization policies or other measures that facilitate economic integration. The implications of these rulings can be observed and evaluated with readily available trade data. This sets the WTO and ECJ apart from international courts of human rights, for example, whose rulings have diverse implications that evade measurement.

Third, I select the WTO and ECJ because they are among the most important international courts. Their rulings bear on integral aspects of international cooperation: how countries may or may not regulate commerce and protect their own domestic economies from global competition. These regulatory issues are of fundamental importance to governments as production and consumption increasingly

cross international borders. Rulings from the WTO and ECJ implicate large sums of money and have consequences for workers, consumers, and thereby politicians who steer national policies accordingly.

Despite the apparent importance of international rulings, their impact on international cooperation is not well understood. This is partly due to methodological reasons. As noted above, it remains extremely difficult to make causal inferences about the effectiveness of international institutions. This dissertation begins to fill the gap in the scholarship by evaluating compliance in economic terms and leveraging appropriate statistical methods.

Among this dissertation’s contributions, I offer a new approach to measuring economic or *de facto* compliance. Because the World Trade Organization and European Union aim to liberalize trade and improve economic integration, they are chiefly concerned with the exchange of goods and services and commerce across borders. My focus on *de facto* compliance reflects this fact.⁴ I use trade data, including bilateral product-level trade flows between disputing countries, to precisely evaluate the impact of international court rulings and infer compliance. This approach complements other recent scholarship that assesses the impact of WTO rulings on trade flows (Bown, 2004*a*; Chaudoin, Kucik, and Pelc, 2013; Bechtel and Sattler, 2015; Bown and Reynolds, 2014, 2015).⁵ Using similar expectations about the role of WTO disputes in liberalizing trade policy, this project specifically emphasizes compliance and heterogeneity in dispute outcomes.

⁴It is also fruitful to examine *de jure* compliance, the policy changes a government adopts in response to international rulings. See Reinhardt (2001).

⁵Bechtel and Sattler (2015) and Bown and Reynolds (2015) both find a positive impact of disputes on trade while Chaudoin, Kucik, and Pelc (2013) find no measurable impact. Bown (2004*a*) explains variation in outcomes focusing on the plaintiff’s threat of retaliation. These studies employ different estimation strategies and levels of aggregation in trade data which may, in part, explain the diverse results. Focusing on adverse rulings and bilateral trade flows between disputing governments, my dissertation in part aims to reconcile these divergent findings by looking at domestic political determinants.

I evaluate compliance by applying the synthetic control method, a new technique for causal inference that is particularly well-suited to the empirical problem at hand. Because *de facto* compliance is fundamentally a matter of whether states conform their behavior to the rulings from international courts, it relies on some notion of what behavior would have been in the absence of those rulings. The empirical difficulty is in estimating such a counterfactual. The synthetic control method is a statistical approach that allows for the construction of an estimated counterfactual based on a weighted combination of multiple control units. It is discussed extensively in Chapter 2.

1.5 Overview of the Dissertation

This dissertation proceeds as follows. Chapter 2 develops a theory of domestic politics and compliance with the rulings of international courts. I test the theory on 125 World Trade Organization disputes using the synthetic control method.⁶ The central empirical challenge is to measure whether countries complied with adverse WTO rulings. Adverse rulings require a defendant government remove barriers to trade. I therefore assess compliance using trade flows. I collect data on product-level trade flows between disputing countries and between the plaintiff and other countries. Using these data and several covariates, I construct a synthetic control that estimates for each dispute a counterfactual trade flow: what trade levels would have been in the post-dispute period absent the legal ruling. I then compare actual trade flows to the counterfactual. The divergence of actual trade from expected levels (synthetic control) yields an estimate of the causal effect of a WTO ruling. Using this procedure, I measure the impact of a WTO ruling on trade for every dispute in my dataset. Over one third of WTO disputes have trade patterns that clearly demonstrate compliance, with many more displaying trade patterns that suggest the defendant removed the

⁶See: Abadie and Gardeazabal 2003; Abadie, Diamond, and Hainmueller 2010, 2014.

illegal trade barrier. With this metric, I then demonstrate that domestic veto points in the defendant government impede compliance. The estimated effect is large and robust. At the WTO, which has limited ability to sanction noncompliance, domestic politics appears to have a pronounced impact on dispute outcomes.

Chapter 3 applies the argument to the European Court of Justice. I show that domestic political constraints in European Union (EU) countries also impact compliance with adverse legal rulings. Here, I focus on infringement disputes over trade-related issues and examine intra-European trade to see whether member states removed barriers to commerce subsequent to the adverse ECJ rulings. One analytical challenge is that all infringement lawsuits are brought by the European Commission, the central bureaucratic body charged with enforcing EU law. Because there is no variation in the plaintiff, I cannot identify trade flows at the country-dyad level. Moreover, because few ECJ disputes concern specific products, I am not able to disaggregate the data at the level required to use the synthetic control method.

To handle this problem, I instead use a hierarchical Bayesian model that captures the multi-level structure of the data. A hierarchical model is appropriate in this context because some variables act on an EU-wide level (e.g. common market rules) and others act on a domestic level (e.g. domestic veto points), each varying over time. Adverse rulings from the ECJ tend to promote intra-EU commerce but the effect is conditional on domestic politics in the defendant government. Defendants with few veto points experience increased intra-EU trade after adverse rulings whereas governments with many veto points do not. The effect is small but statistically significant. Even the European Court of Justice, a relatively strong court with the ability to sanction noncompliance, is sensitive to domestic politics.

Chapter 4 extends the theory to show how institutional design can mitigate the link between domestic politics and international cooperation. Comparing the WTO Dispute Settlement Mechanism to the European Court of Justice, I argue these international courts share many design features, but differ significantly in their flexi-

bility. The WTO Dispute Settlement Mechanism is more flexible, imposing low costs for noncompliance and allowing countries more opportunity to negotiate and delay implementation of adverse rulings. The European Court of Justice is less flexible, imposing high costs for noncompliance, enforcing strict time-lines for dispute settlement and providing governments little opportunity to negotiate settlements. I argue these institutional differences mitigate the impact of domestic politics on dispute resolution and provide empirical evidence.

This chapter demonstrates that legal disputes take longer to resolve when the defendant country has many veto players. The dispute duration includes the litigation period and, in the case of an adverse ruling, the period in which the defendant country implements the court's verdict. Typically, disputes are prolonged when the defendant government resists and retains non-compliant policies. I show disputes consistently last longer when the defendant government is highly constrained. This tendency is pronounced in WTO disputes, where delays and noncompliance carry only modest penalties. At the ECJ, which imposes strict time-lines and punishes defiant governments, this tendency is not significant. The design of the institution appears to mitigate the impact of domestic politics on dispute resolution. Dispute duration is thus another important way to gauge the successes and failures of international courts.

Together, these chapters show that international courts can change economic outcomes but their effectiveness hinges on domestic politics and institutional design. Chapter 5 concludes by highlighting the main contributions of this dissertation, raising additional questions about the role of courts in the international economy, and pointing to broader implications for the optimal design international institutions.

CHAPTER 2

When are International Institutions Effective?

The Impact of Domestic Veto Players on Compliance with WTO Rulings

2.1 Introduction

The World Trade Organization (WTO) Dispute Settlement Mechanism, created in 1995, promised a new institutional solution for trade disputes between countries and an advance in international economic cooperation. Compared to its predecessor, the General Agreement on Tariffs and Trade (GATT), the Dispute Settlement Mechanism provides a standardized process through which lawsuits are conducted, verdicts are delivered, and implementation is monitored. Nearly all WTO rulings identify infringements and require the respondent government modify its policies. Yet only some governments comply readily; others delay for years or defy the WTO altogether. What explains this variation?

For example, in 1995, the European Union (EU) sued Japan over an alcohol tax that favored domestic producers and inflated the price of comparable foreign products. Two years later, the EU sued Korea on nearly the same grounds. In both cases, the WTO ruled against the respondent governments, ordering them to modify their tax codes. The Korean government, with its recently consolidated political power, promptly complied. But Japan, with its domestic partisan discord, did not. Do features of these countries' domestic politics explain the divergent outcomes?

Many scholars argue that democracy facilitates international economic cooperation (e.g. Martin 2000; Mansfield, Milner, and Rosendorff 2002). Institutional divisions of power and partisan opposition in government—hallmarks of a functional democracy—create multiple veto points (Tsebelis, 1995). Multiple veto points narrow the set of international agreements that can be ratified, since more domestic actors must coordinate and consent to the agreement (Milner and Rosendorff, 1997).¹ This makes the treaties enacted by democracies more credible commitments about governments’ future behavior than those by non-democratic regimes. Veto players can lock in cooperative policies and make it difficult for governments to renege on their international promises.² However, this lock-in mechanism has a second side. Contrary to a perspective prevalent in the literature, I argue veto players can sometimes hinder international cooperation.

Democratic leaders can and do break international commitments when the domestic pressure to do so is acute. For example, a government might grant farmers’ request for import restrictions amid sharply declining wheat prices. Or, faced with a forthcoming election, a leader might introduce a subsidy that curries favor from key interest groups. Many scholars argue that institutions that allow leaders to respond to such pressure by temporarily violating their commitments are more stable (e.g. Staiger and Bagwell 1999; Rosendorff and Milner 2001; Bown 2002; Rosendorff 2005). By permitting some violations, flexible institutions allow members to manage temporary economic shocks or political circumstances without abandoning the institution altogether.

However, just as multiple veto players can lock in international commitments, they can also lock in violations of those commitments. Once an initial violation

¹Some studies, e.g. Mansfield, Milner, and Pevehouse (2007), argue that veto players reduce the probability of forming a trade agreement but do not address whether the governments follow through on their commitments.

²And democratic leaders, beholden to voting publics, might be wary about breaking international agreements, lest they generate audience costs (Tomz, 2007; Dai, 2005).

occurs, a government with more veto players is less likely to return to cooperation. For these cases, flexibility mechanisms may be less successful in generating long-term stability. In WTO disputes, a losing respondent government with many veto players is less likely to comply with the legal ruling and correct the initial breach. Defining features of democratic politics can actually obstruct international economic cooperation.

I evaluate government compliance with adverse WTO rulings by estimating the causal effect of the rulings on product-level trade flows between the disputing countries. True causal effects are difficult to identify because they rely on specifying the correct counterfactual (Rubin, 2005). Using the method of synthetic case control (Abadie and Gardeazabal, 2003; Abadie, Diamond, and Hainmueller, 2010, 2014), I estimate a plausible and precise counterfactual that represents what product-level trade would have been without the WTO's legal verdict. For every dispute, I identify the affected products and examine trade flows of just those products. Then by comparing actual trade to the estimated counterfactual, I determine whether or not trade flows in the disputed products increased in relative terms, beyond ordinary fluctuations. Relative deviations in trade reveal the approximate causal effect of an adverse WTO ruling, with a positive deviation indicating the respondent government complied. I measure compliance in all 125 WTO disputes between 1995 and 2011 that received an adverse ruling on an import-restricting trade policy.

Using this novel metric, I show compliance patterns at the WTO reflect the cross-cutting impact of democratic politics. Some rulings from the WTO Dispute Settlement Mechanism (DSM) generate marked increases in trade indicative of compliance while others have no detectable effect. The variation is systematically linked to domestic politics. As the number of veto players in the respondent government increases, compliance with adverse rulings becomes less likely. Veto players appear to lock in initial violations and obstruct compliance, even when international pressure to capitulate is substantial. The result is large and robust, controlling for other

potentially confounding factors. The size of the complainant's and respondents' economies do not reliably predict compliance, nor does the extent of the adverse legal ruling. While prominent and influential, the DSM is not insulated from the domestic political controversies that lead to trade disputes in the first place.

Complementing the institutional design literature, this paper offers an empirically rigorous appraisal of institutional effects. Some scholars have puzzled over the seemingly insignificant impact of WTO disputes on trade (Chaudoin, Kucik, and Pelc, 2013). Measuring compliance through bilateral, product-level trade flows, I show there is substantial variation in dispute outcomes that is obscured when one examines the compliance record on average. This heterogeneity suggests that international dispute settlement may not be uniformly effective or ineffective. Scholars continue to disagree about the impact of international institutions on cooperation (e.g. Mearsheimer 1994; Goldstein and Martin 2000; Hafner-Burton, von Stein, and Gartzke 2008; Voeten 2013). Inferential obstacles are partly to blame. When states appear to comply—for example by removing trade barriers—are they doing so in response to their international commitments? Or are they implementing policies they would have selected anyway? Because countries voluntarily enter international agreements they intend to follow, it is little surprise the compliance record *appears* to be good (Downs, Roake, and Barsoom, 1996; Von Stein, 2005). I weigh in on the debate by measuring *de facto* compliance in a precisely defined and particularly revealing domain: WTO disputes with adverse rulings. Construed this way, my evidence casts serious doubt on the hypothesis that WTO rulings have no effect.

The following section situates my study in existing literature and describes dispute settlement at the WTO. The third section presents a theory of domestic politics and compliance with international institutions. Fourth, I discuss obstacles to assessing compliance and my analytic solution. In the final sections, I present statistical results and discuss the findings.

2.2 Cooperation, Compliance, and the WTO

International institutions aim to improve cooperation among countries by solving collective action problems (e.g. Keohane 1982; Stein 1990). In international trade, countries have a unilateral incentive to impose trade barriers but a mutual interest in exchanging goods and services freely. Trade agreements aim to move from a sub-optimal outcome of high trade barriers to a more efficient one. Yet international institutions struggle with compliance. Because they represent voluntary agreements among countries, they have difficulty ensuring governments abide by treaty terms. Sometimes noncompliance is the result of incomplete information. For example, a violation may reflect ambiguous or imprecise legal obligations (Chayes and Chayes, 1993). Other times, governments are aware of obligations but behave opportunistically. Violations often go unnoticed and enforcement is informal and decentralized. Governments may choose to violate whenever the benefits of doing so outweigh the expected penalties. Compliance problems have broad effects. For example, when governments violate trade agreements, they undermine the collaborative equilibrium: other states then prefer to impose trade protection. Noncompliance by a few leads to a worse outcome for many.

States often design international institutions with flexibility mechanisms in order to mitigate compliance problems (Abbott and Snidal, 1998; Koremenos, Lipson, and Snidal, 2001; Carrubba, 2005). Flexibility mechanisms allow governments to sometimes violate treaty obligations without abandoning the institution altogether. By permitting short-term violations under certain circumstances, flexibility mechanisms can enhance long-term compliance if they prevent countries from leaving international agreements (e.g. Staiger and Bagwell 1999; Bown 2002). Dispute settlement mechanisms are institutional features that some scholars believe increase flexibility and thereby improve stability (e.g. Rosendorff 2005; Kucik and Reinhardt 2008; Pelc 2009).

Scholars have identified two forms of compliance in the context of international relations (Fisher, 1981; Simmons, 1998). First-order compliance is when a state abides by its substantive legal obligations. In the WTO, a government complies in the first sense if its trade policy and practice reflect substantive legal commitments, e.g. low trade barriers. When a state fails to comply with treaty terms and a dispute arises, it is often required to correct the violation. Second-order compliance is when the state adjusts its policy and practice after an initial violation, often in response to international litigation. A government complies in the second sense if it loses a dispute and corrects the violation, e.g. removes an illegal trade barrier.

Governments that have been harmed by the actions of another state often use international dispute settlement mechanisms to address their grievances, especially when there are domestic political benefits from doing so (Goldstein and Steinberg 2008; Davis 2012). Some governments seek legal recourse for treaty violations in order to capture the support of key domestic constituencies or as political cover for unpopular policies (Allee and Huth, 2006). Domestic politics factors into a government’s decision to *initiate* litigation. It should also influence a government’s *response* when it is convicted of violating international commitments. Many studies focus on dispute initiation; few consider compliance *ex post*. The research that does largely ignores domestic politics, instead emphasizing features of the international system (e.g. Bown 2004a).

The WTO is an ideal venue to examine compliance. First, it is the cornerstone of the multilateral trade regime. If any international institution is to have a demonstrable effect on its members’ behavior, the WTO is it (Rose, 2004; Goldstein, Rivers, and Tomz, 2007; Subramanian and Wei, 2007).³ Second, its Dispute Settlement Mechanism is seen as the exemplar for international dispute resolution. Many studies appraise the DSM’s design and function (Hudec, 1999; Busch and Reinhardt,

³Not all studies find a positive impact on trade and some find conditional effects, e.g. Gowa and Kim (2005). However all agree the WTO is a critical venue.

2000). With widespread interest among political scientists, lawyers, and economists alike, there is an implicit assumption that WTO legal verdicts matter. Yet scholars are only beginning to collect hard evidence of whether the rulings restore trade.

The primary objective of the WTO is to promote trade among countries by constraining the trade policies of member states. WTO disputes arise when one government alleges another has imposed trade barriers that violate treaty terms. Trade barriers differ in form. For example, subsidies lower production costs for domestic industries giving them a competitive advantage and reducing demand for imported goods. Tariffs inflate the price of imports, again improving domestic industry's competitive advantage. Quantitative restrictions place a limit on the volume of imported products. Regardless of form, these policies effectively limit imports into the respondent country.⁴

The DSM aims to restore trade cooperation. As the WTO itself states, “the priority is for the losing defendant to bring its policy into line with the ruling.”⁵ This distinguishes the DSM from other institutions which aim to punish violators, levy fines, etc. Adverse rulings require the respondent remove trade barriers, prompting an increase in imports. These trade flows can be observed and measured, making the WTO an excellent setting to study compliance.

Moreover, WTO disputes lend special insight into institutional effectiveness. As Downs, Rocke, and Barsoom (1996) observed, “we do not know what a high compliance rate really implies. Does it mean that even in the absence of enforcement states will comply with any agreement...or does it mean that states only make agreements that do not require much enforcement?” (383). Ordinarily, it is problematic to infer that an institution is effective because states comply with its rules (Martin, 2013). But by focusing on situations where states do not want to cooperate and nevertheless

⁴For a discussion of trade barriers in GATT/WTO disputes, see: Staiger and Bagwell (1999).

⁵“What is the WTO? Understanding the WTO,” available: <http://www.wto.org>.

do—second-order compliance—we *can* draw inferences about effectiveness. This is what makes WTO disputes so revealing. By virtue of being sued and found guilty of violations, we know respondent government’s preference is to not comply. When the respondent corrects the violations in spite of this preference, we can infer the influence of the institution.

Besides the disputants, other states often have a stake in WTO disputes. The respondent’s trade policy might harm them in a similar way as it does the complainant. Or the case may establish standards akin to legal precedent (Pelc, 2014). Since WTO disputes have broader importance than their effect on the complainant and respondent, the DSM allows states to enter the proceedings as “third parties”. Third party governments usually support the complainant. Their presence reflects the perceived international importance of the dispute. They may generate international pressure to comply by increasing the publicity of the dispute, monitoring to ensure the ruling is implemented faithfully, and making enforcement more likely.

WTO disputes have several stages. The first is formal negotiation. At least 60% of disputes have been settled early through bilateral agreements. When early settlement is impossible, the dispute moves into litigation. A panel evaluates the case on the merits and issues a ruling called the “panel report.” While rulings consist of many separate claims, they usually require the respondent change some aspect of its trade policy. The vast majority (94%) favor the complainant on at least one legal claim. Provided the ruling is not changed through appeal, the respondent government is given a “reasonable period of time” to correct the violation. This is the compliance deadline. If a respondent fails to make required policy changes before the deadline, it has not complied.

The WTO has limited enforcement options. If the implementation deadline passes and the losing respondent fails to take adequate action, the complainant may initiate compliance proceedings. Complainants can only punish noncompliance after all other

legal remedies are exhausted.⁶ Retrospective penalties are prohibited. As a result, respondent governments have an incentive to prolong the dispute and delay compliance. By delaying, they reap the benefits of noncompliance—i.e. advantages from unilateral trade barriers—without incurring costs, creating a “remedy gap” (Brewster, 2011). When do governments implement adverse WTO rulings in a timely manner? When do they fail to comply?

2.3 A Theory of Domestic Veto Players and Compliance

2.3.1 Domestic Preferences, Power, and Institutions

Within a country, different political actors have divergent preferences over international affairs. Some prefer trade protection; others trade liberalization. Many interest groups represent import-competing industries and prefer protectionist policies. Trade liberalization poses a threat to these groups by expanding imports, increasing competition, and driving down prices of the goods they produce. Voters sometimes have preferences over trade policy, but compared to interest groups, their positions are not well-formed or informed (Hiscox, 2006).⁷ As consumers, voters may prefer the availability of lower-cost goods that trade liberalization brings but they face well-known problems in mobilizing. Interest groups and voters impose political pressure on politicians.

Politicians vary in their own inherent preferences as well as in the political pressure they experience. Politicians in different branches of government experience domestic pressure in various ways, depending on the composition of their constituencies. Legislators often prefer more protectionist trade policies than prime ministers

⁶The WTO can authorize the complainant to retaliate through counter-measures.

⁷Some research suggests that voters do not understand their own economic interests and routinely support policies that harm them. This tendency may be particularly pronounced for trade policy (Kono, 2006).

or presidents do (Lohmann and O'Halloran, 1994). Legislators respond to local constituencies who benefit from protection for specific industries.⁸ By contrast, executives, with their broad constituencies, are more sensitive to the aggregate benefits of international trade cooperation.

Exogenous economic and political shocks can change the relative power of different interest groups and voters. An economic downturn that acutely harms an industry may lead groups to pressure their politicians for policies that benefit them. Typically, these groups are highly informed about whether current trade policy is helping or hurting their business. Subjected to increased pressure from mobilized industry groups, politicians often shift their preferences. Similarly, political shocks can shift different groups' relative power. Political responsiveness often changes for reasons unrelated to trade policy. For example, an upcoming election can make politicians sensitive to preferences in a key electoral district. Under these conditions, politicians are more likely to implement trade protection policies that provoke international disputes.⁹

Countries vary in the degree to which political authority is concentrated within government. Some governments have substantial divisions in authority. In these governments, it is harder to change existing policies because more political actors can block change. In part, the divisions arise from domestic political institutions. Checks and balances make policy change difficult. Divisions also arise from party politics. The partisan composition of the government can make it more or less difficult to change policies. When the legislature features a relatively strong opposition party or many opposition parties, there are significant obstacles to policy change. Taken together, domestic institutional and partisan divisions form *veto points* in govern-

⁸Export-oriented industries sometimes push policy in the opposite direction (Gilligan, 1997).

⁹For example, when the Canadian magazine industry suffered unprecedented decline, industry officials urged parliament to impose a controversial tax to stave off foreign competition. See: "Magazine Industry Urges Laws to Stem Foreign Competition" *The Toronto Star*, February 9, 1993.

ment. For example, if one branch of government is held by a first party and another branch is dominated by a second party, the government has especially divisive domestic politics—many veto points. As veto points increase, political authority becomes more fragmented and policy becomes more difficult to change (Tsebelis, 1995; Henisz, 2000; Tsebelis, 2003).

Veto points determine how responsive governments are to shifts in preferences and power. When a government has few veto points, only a few political actors must coordinate and it is relatively easy to change policies. So minor shifts in industry preferences or political power can have a potent influence on policy. Conversely, when government has many veto points, policy change is quite difficult because it requires coordination among many political actors. Only major shifts in preferences or power can alter policy.

2.3.2 International Cooperation and Compliance

Domestic veto players constrain governments when they join international agreements. Divisions make it more difficult for a government to join an international agreement because “[m]ultiple veto players...narrow the set of [treaty] proposals that can be domestically ratified” (Simmons, 2009, p. 69). When domestic government becomes more divided, treaties are less likely to be ratified because there are more domestic actors willing to block the agreement. Domestic veto players can limit international cooperation in many domains. Because trade preferences of different domestic groups often vary widely, trade agreements are particularly susceptible to such divisions. Divided domestic politics tends to constrain trade liberalization (Milner and Rosendorff, 1996, 1997).

After a country has joined an international trade agreement, veto players are thought to facilitate compliance with the agreement. Governments with substantial divisions in authority may be less responsive to political pressure and industry

demands for new trade protection. In these governments, interest groups will have to persuade many political actors to support a new policy. Interest groups often pressure politicians for trade protection that benefits their targeted industry, sector, or region. Sometimes the protection they demand violates trade agreements. Veto players make it more difficult for a government to reverse existing trade policies. Following this logic, trade agreements formed by democratic governments are thought to represent more credible commitments about future trade policy (Mansfield, Milner, and Rosendorff, 2002). By locking cooperative policies in place, veto players should improve a government's first-order compliance.

However, empirical obstacles make it difficult to test the impact of domestic politics on first-order compliance. First, it is difficult to observe when a leader is tempted to violate international commitments because domestic political pressure is idiosyncratic. Second, governments (and researchers) may be uncertain about what constitutes a violation of treaty terms. WTO disputes often focus on whether particular trade measures are legal or illegal; the answer is often nonobvious. Third, even if one could perfectly differentiate compliant trade policies from violations, not all trade violations are observed—only those which trigger a dispute.¹⁰ Governments impose trade barriers that go uncontested and these violations are hard to identify. As a result, an empirical analysis of first-order compliance may generate biased conclusions about the impact of domestic veto players.

Dispute settlement mechanisms make second-order compliance easier to examine than first-order compliance. International legal disputes reveal information about the initial violation and resolution. Although domestic political pressure is idiosyncratic, litigation publicizes these interests, revealing the domestic factors that lead to the violation. Litigation generates rulings that highlight infractions and explicitly

¹⁰Some studies have examined governments' choices to impose legal safeguards versus illegal protection that provokes a dispute (Bown, 2004*b*). These measurable instances represent a small subset of cases.

differentiate legal from illegal trade measures. And the respondent government that loses a dispute is forced to respond. Some will adjust their policies and behavior to implement the legal ruling; others will not. Second order compliance and defiance are both observable. Finally, by drawing attention to government behavior and specifying the requirements for implementation, dispute settlement mechanisms provide criteria against which one can evaluate second-order compliance.¹¹

Just as veto players constrain a government's ability to form and violate trade agreements, they make it more difficult to restore cooperation, after a violation has occurred. When a government fails to comply with its substantive international commitments, domestic political divisions can lock in the violation. Sometimes, reversing the violation may require nullifying an old law or passing a new one. This is common for regulatory trade protection. Political actors who benefit from the trade barrier have incentive to block policy change. Veto players in domestic government create points where policy change can be obstructed. For example, some legislators may block legislation that implements a WTO ruling.¹² There will be some instances where policies can be changed by unilateral acts, like executive orders. But on average, second-order compliance becomes less likely as domestic veto points increase.

Faced with an adverse ruling from the WTO Dispute Settlement Mechanism, leaders decide whether and when to second-order comply. This decision is shaped by both short-term economic conditions and long-term institutional constraints. Leaders who face acute exogenous shocks have little short-term *incentive* to comply with rulings. Leaders who face more domestic veto points experience greater institutional

¹¹For example, a dispute between the US and Canada (DS31: Canada-Periodicals) highlighted political pressure from Canada's waning publishing industry. The WTO ruled that Canada's magazine tax violated treaty terms, establishing fault. And the ruling laid out requirements for compliance.

¹²In the magazine dispute, partisan fighting between Canada's Tory and Liberal parties and between the upper and lower houses of Parliament obstructed compliance with the WTO ruling. See: "Magazine Bill Angers Senate; Tories Say Liberals Trying to Shove Legislation Through," *The Hamilton Spectator* (Ontario, Canada), June 1, 1999.

and partisan political *obstacles* to complying. Conditional on going to trial and losing the case, a leader who faces more veto points at home has an incentive to wait for the exogenous shock to pass. Rather than complying immediately, a leader who faces more veto points, and thus high obstacles to compliance, is more likely to violate the agreement until the political or economic conditions change. *Conditional on violating a treaty, domestic political divisions (veto points) should be associated with less compliance.*

Domestic veto players are not the only factor driving second-order (non)compliance. A country that loses a WTO dispute and does not change its policy before the implementation deadline may be subject to enforcement. To isolate the impact of domestic veto players, one must control for international pressure. International pressure can improve the prospects for compliance. Litigation publicizes the initial violation, raising awareness among other countries. Other countries can reduce the respondent's benefits from noncompliance by denying future opportunities for cooperation or can impose retaliatory trade measures that increase the respondent's costs from noncompliance. International pressure may partly offset the effect of domestic pressure.

2.3.3 Illustrations

Two examples illustrate ways in which domestic political divisions—both institutional and partisan—can hinder compliance.¹³ In the first case, the respondent government had few divisions and complied quickly. In the second case, the respondent had more divisions and defied the WTO ruling for many years before eventually capitulating.

¹³The Japanese and Korean domestic political circumstances I describe are also mentioned in Mansfield and Milner (2012), who show that veto players impact a country's ability to form preferential trade agreements.

In the first case, the European Union sued South Korea over its liquor tax system (DS 75: Korea - Taxes on Alcoholic Beverages, 1997). European whiskey sales to Korea had fallen sharply and EU officials argued this was because the Korean government violated its WTO obligations by levying a 130% tax on alcohol imports but not on the local product, soju.¹⁴ The WTO determined that the tax constituted a trade barrier and ordered Korea to revise its system.¹⁵ The Korean government was under political pressure to defy the WTO. As one newspaper reported, “at stake besides soju itself are votes....soju is the drink of choice for South Korea’s poor, a group the government is wooing before next April’s parliamentary elections.”¹⁶ But at that time, Korea had relatively few veto points. The Grand National Party had recently lost a major election and the opposition party, the National Congress for New Politics, consolidated its power. With a unicameral legislature, only modest partisan divisions, and thus few veto points, Korea promptly revised its tax law.¹⁷ Korea had minor domestic obstacles and complied with the adverse ruling.

In the second case, the EU sued Japan on similar grounds (DS 8: Japan - Taxes on Alcoholic Beverages, 1995). Reviving an unresolved dispute under the GATT,¹⁸ the EU argued the Japanese tax system was discriminatory because it levied a substantially higher tax on foreign products than the local alcohol, shochu. The WTO ruled against Japan, ordering it to reform. At the time, Japan had many veto points because the long-ruling Liberal Democratic Party had been forced to form a coalition government with its competitors. With substantial domestic obstacles, Japan did not

¹⁴ “Commission Calls for WTO Talks on South Korean Alcohol Tax,” *European Report*, April 2, 1997.

¹⁵ “Tax ruling boosts whiskey hopes,” *The Herald*, Glasgow Scotland, January 19, 1999.

¹⁶ “S. Korea’s most popular drink under fire: Government must raise excise tax on soju and cut import tariffs on whiskey” *The Vancouver Sun*, British Columbia, October 23, 1999.

¹⁷ “EU, S. Korea Becoming Closer, More Interdependent,” *The Korea Herald*, October 16, 2000.

¹⁸ “Japan - Customs Duties, Taxes, and Labeling Practices on Imported Wines and Alcoholic Beverages,” GATT Report of the Panel, November 10, 1987 (L/6216).

comply. Industry leaders in the EU attacked the Japanese government for “dragging its feet on the implementation of [the] WTO ruling...over as long a period as they think they can get away with.”¹⁹ Despite a compliance deadline of February 1998, Japan decided to “gradually increase the tax on shochu...[through] October 2001.”²⁰ With a bicameral legislature, divided domestic politics, and thus many veto points, Japan defied the WTO ruling for many years.²¹

These examples illustrate the variation in outcomes. When the EU won the lawsuit against Korea (few veto points), the Korean government complied in a timely manner. Yet when the EU won the lawsuit against Japan (many veto points) there was a prolonged period of noncompliance. The difference between Korea and Japan’s veto points is significant.²² As domestic government grows more divided through institutional and partisan divisions, the prospects for compliance wane. While the examples provide preliminary support for the theory, a more systematic test requires evaluating *de facto* compliance in all WTO disputes with adverse rulings. I develop a method for measuring compliance using the most objective metric available: product-level trade flows. I then demonstrate that conditional on an adverse ruling, domestic divisions obstruct a return to compliance on average.²³

¹⁹“Japan fails to remove whiskey tax,” *The Scotsman*, Scotland, November 23, 1996.

²⁰“WTO Ruling Pushes Shochu Makers to Reinvent Product” *The Nikkei Weekly*, Japan, March 31, 1997.

²¹Ultimately, the disputants reached a “mutually acceptable solution” that included short-term compensation and long-term policy reform. “Japan - Taxes on Alcoholic Beverages: Arbitration under Article 21(3),” WTO Document No. 97-0558, February 14, 1997.

²²Veto points are measured on a scale of 0 to 1 with larger values denoting more divisions. At the time of the lawsuits, Korea had 0.45 veto points and Japan had 0.60. The difference (0.15) exceeds one standard deviation in my sample (0.12) and is equivalent to the difference between Israel, with a unicameral parliament, and Italy, with a bicameral parliament (Henisz, 2000).

²³There may be multiple ways that governments can alter their policies to comply. For example, in the Korean alcohol dispute, Korea could have complied by lowering the tax on foreign liquor, raising the tax on local soju, or a combination (actual outcome). Some governments may chose compliance strategies that activate fewer veto players. These cases should bias my results toward a null finding.

2.4 Assessing Compliance with WTO Rulings

2.4.1 Measurement Strategy

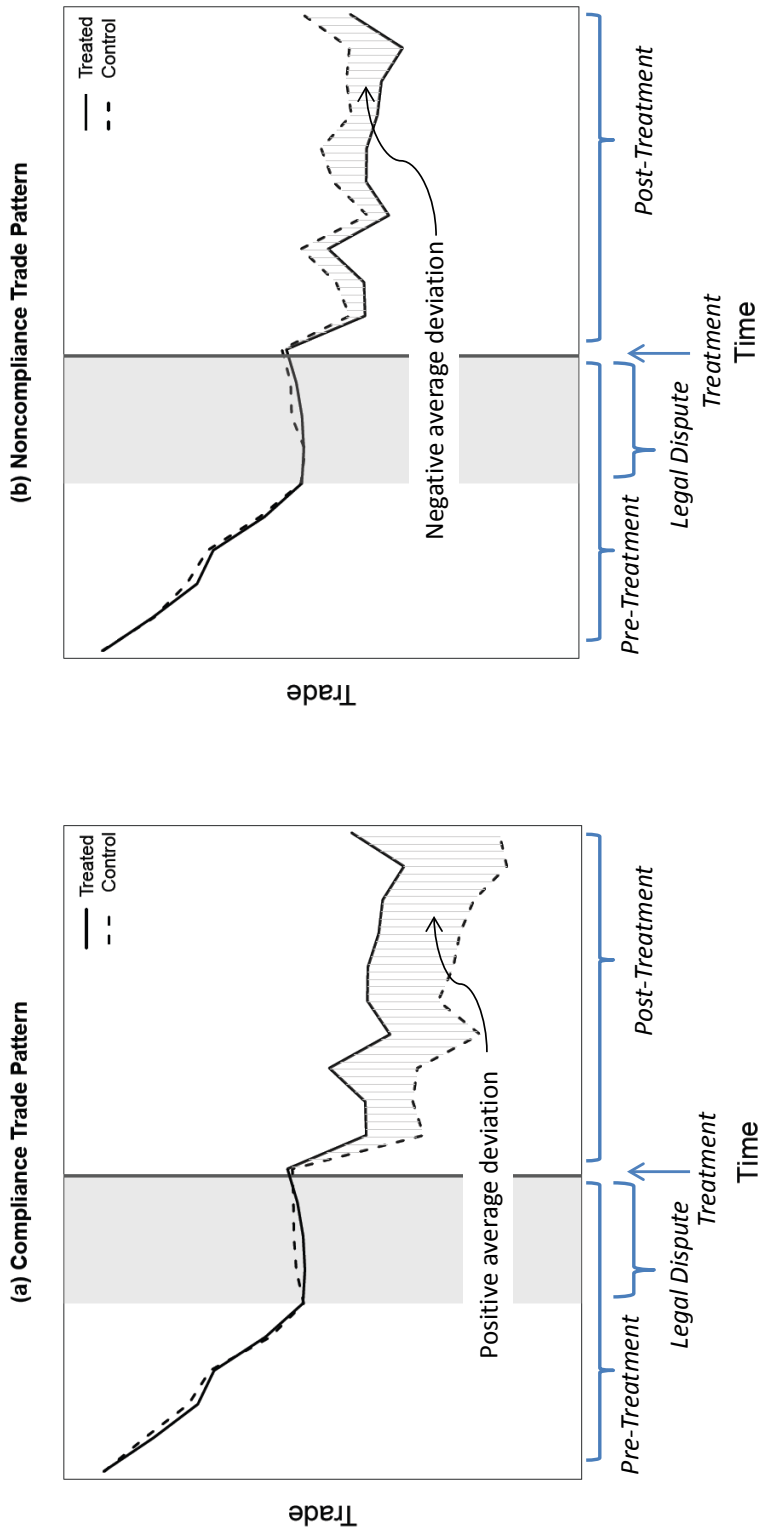
An ideal test of the DSM’s effectiveness would compare the outcome with litigation to the outcome of the same dispute where the DSM did not exist. Because nearly all countries are members of the WTO, it is impossible to construct such a counterfactual. Yet one can still examine effectiveness with respect to particular aspects of the institution. Adverse WTO rulings require a respondent government remove an impermissible trade barrier. If the trade barrier is removed, imports into the respondent country should subsequently increase. The empirical challenge is to correctly identify such an increase.

Absent a randomized experiment, the only way to estimate the causal effect of an adverse WTO ruling is by comparing trade between disputing countries to an estimated control that represents what trade would have been without the adverse ruling. This paper uses observational data to construct an approximate control in order to estimate the effect of the adverse ruling on trade and thereby assess compliance. Although adverse rulings are not randomly assigned and true treatment effects cannot be obtained, casting the analysis in a causal inference framework is informative. I adopt the language of experiments and refer to the adverse ruling as the “*treatment*” and post-ruling trade as the “*outcome*.”

One obstacle is that trade often changes for reasons unrelated to the WTO dispute itself. WTO disputes may be prompted by economic trends that cannot be reversed, even when the respondent government complies. These factors may be confounded with the effects of the WTO dispute. For example, governments sometimes initiate disputes when their exports for a product are already declining, even though the decline is partly driven by forces exogenous to the respondent’s trade barrier. My technique accounts for this.

Figure 2.1 illustrates two hypothetical scenarios for WTO disputes. In both, trade is decreasing over time and a WTO dispute yields an adverse ruling (treatment). Observed trade (outcome for the treated unit) continues to decrease after the ruling (during the post-treatment period). Did the respondent government comply?

Figure 2.1: Hypothetical Compliance and Noncompliance Trade Patterns



Note: Hypothetical scenarios for a trade dispute where trade is trending down over time. Panel (a) illustrates a trade pattern showing compliance, indicated by a positive average deviation. Panel (b) illustrates a trade pattern showing noncompliance, indicated by a negative average deviation. The shaded area denotes the duration of the legal dispute and the vertical line indicates the implementation deadline.

This depends on what trade *would have been* after the WTO dispute, if there were no adverse ruling (outcome for the control unit). When trade for the treated unit exceeds trade for the control unit, as measured by a positive average yearly deviation after the treatment, I infer the respondent complied (Figure 2.1(a)). Conversely, when trade for the treated unit does not exceed trade for the control unit, I expect the respondent did not comply. Because trade data are noisy, I aim to avoid “false negatives” and infer noncompliance only when I detect a negative average yearly deviation, (Figure 2.1(b)). To estimate the counterfactual, I create an approximation of a control unit from a statistically optimal combination of other countries’ trade, as detailed below.

This trade-based, *de facto* measure of compliance prioritizes economic outcomes.²⁴ Of course, like any metric, it is imperfect. Governments sometimes settle WTO disputes through compensation schemes where the losing respondent offers a payment or trade concession to the complainant in lieu of prompt implementation,²⁵ and these cases will not count as compliance per my methodology. But since the central goal of the WTO is to liberalize trade, my trade-based measure best reflects this goal.

2.4.2 Product-Level Trade Flows

To implement the measurement strategy, I select the 125 WTO disputes that (1) received an adverse ruling between 1995 and 2011 and (2) concerned import restrictions.²⁶ The sample consists of all disputes with a panel ruling that favored the complainant on at least one legal claim. I exclude disputes where the ruling was completely overturned on appeal. I include only disputes about import restrictions

²⁴By contrast, a policy-based, *de jure* measure of compliance emphasizes legal outcomes.

²⁵Dispute Settlement Understanding Article 22.1

²⁶The data and technical appendix will be available upon publication on my website.

like tariffs, countervailing duties, anti-dumping measures, safeguards, quantitative restrictions, discriminatory tax schemes, or other barriers to trade.²⁷

WTO disputes cite specific products and services. For every dispute, I identify these products and collect annual bilateral trade data for each of them, aggregating when multiple products are cited in a given dispute. Products are classified by the Harmonized System (HS) codes and services by the Central Product Classification codes. Only the trade flows for disputed products and services enter the analysis.²⁸ These data come from the UN Commodity Trade Statistics Database, the UN Service Trade Statistics Database, and the European Commission’s Eurostat database on international trade.

The unit of analysis is the directed dyad-year. For each dispute, I use the complainant’s annual exports of that disputed product to the respondent. To build the counterfactual “control unit,” I use the complainant’s annual exports of the disputed product to other countries not engaged in the dispute. In each dispute, I choose up to fifteen other countries to estimate a synthetic control unit, as discussed below.

Trade is measured as the “export share,” the complainant’s annual exports of disputed products to the respondent or other country, divided by its total annual exports of the products to the world. Because some disputes concern products with a large value (e.g. gasoline) and others with a small value (e.g. preserved peaches), the value of exports varies widely from one WTO dispute to the next. The export share standardizes trade values to improve comparability across disputes. It also controls for price fluctuations and variation over time in the complainant’s export volumes. A large export share means the respondent’s market was very important

²⁷All disputes are categorized as either import-restriction cases or export-promotion cases. In the online technical appendix, I include in the sample the 24 disputes over export-promoting measures like subsidies. The findings hold for the larger sample.

²⁸Wherever possible, the six-digit level of HS codes are used. If disputes cite products at the four- or two-digit level or have insufficient coverage, I use the highest level of precision available.

to the complainant.²⁹ For each dispute and in each year t , the complainant exports disputed products to the respondent and to other countries. Let $j = 1$ denote the respondent and let $j = 2, 3, \dots, J$ denote the other countries. Then country j 's export share of the products in year t is:

$$\text{Export Share}_{jt} = \frac{\text{Complainant's Exports of Product to Country } j_t}{\text{Complainant's Exports of Product to World}_t}.$$

Alternative model specifications using export values—rather than export shares—confirm the main results and are presented in the online technical appendix.

To satisfy fundamental assumptions for causal inference, I transform the export share. The “stable unit treatment value assumption” requires that the treatment status of one unit does not affect the potential outcomes of the other units—non-interference between units. Yet export shares are compositional data: when the respondent's export share increases, other countries' export shares necessarily decrease. This is a problem because the treated unit and control units will vary in inverse proportion to one another. A common way to deal with dependence across compositional units is to convert to log ratios (Tomz, Tucker, and Wittenberg, 2002). This factors out the proportional component and isolates the independent variation among the units. In each dispute, I divide by the *ex ante* largest trade partner among control countries. The transformed unit of analysis is:

$$y_{jt} = \log \left(\frac{\text{Export Share}_{jt}}{\text{Export Share}_{2t}} \right), \quad \text{for } j \in \{1, 2, 3, \dots, J\}$$

²⁹Export shares improve comparisons between countries and within countries over time. Another alternative is to standardize relative to the respondent's imports but this may generate biased results because in many WTO disputes, the respondent imposes trade a barrier against all imports, regardless of the country of origin.

where $j = 2$ denotes the control country (in $j = 2, 3, \dots, J$) with the largest export share.

This transformation mitigates the potential bias that can arise from trade diversion. Trade diversion, the increase (reduction) in trade some countries may experience when another country imposes (removes) trade barriers, violates the stable unit treatment value assumption. Dividing by the export share for the largest trade partner and taking the logarithm controls for the largest expected amount of trade diversion.³⁰ This isolates fluctuations in trade that are above and beyond trade diversion.

Countries other than the disputants often enter WTO disputes as third parties, but this is not a serious source of bias in my measurements either. Like the complainant, third parties are typically concerned with their exports to the respondent (Bown, 2005). When the respondent complies, third parties should benefit in the same way as the complainant does: their exports to the respondent increase.³¹ So the presence of third parties is unlikely to generate systematic bias. I estimate the counterfactual from the complainant's exports to other countries; *not* other countries' exports to the respondent.³²

2.4.3 Synthetic Control Method to Measure Compliance

I use the synthetic control method to estimate the counterfactual (synthetic control unit) and infer the approximate causal effect of an adverse WTO ruling (Abadie and

³⁰This is most effective when trade is diverted evenly to all countries in the donor pool. There is also a possibility of substitution effects among different products.

³¹Consistent with a rational choice framework, I assume complainant governments only file WTO complaints when their expected benefits exceed the significant legal costs. This implies that complainants tend to be the primary beneficiaries of WTO lawsuits; else a different country would have sued.

³²See the Appendix.

Gardeazabal, 2003; Abadie, Diamond, and Hainmueller, 2010).³³ When the units of analysis are a few aggregate entities like countries, a combination of comparison units often does a better job reproducing the characteristics of the unit than any single comparison unit alone (Abadie, Diamond, and Hainmueller, 2014). The control unit in the synthetic control method (SCM) is constructed from a weighted average of all potential comparison units. To implement the methodology, I use the ‘‘`synth`’’ package in R (Abadie, Diamond, and Hainmueller, 2011).

For every dispute, I use a sample of countries observed over multiple years. The respondent country is the “treated unit” because it is “treated” with an adverse ruling. Other countries form the “donor pool,” the potential comparison units that are “untreated” and used to approximate the counterfactual. I manually select these countries for each dispute to provide the best possible match and exclude sources of bias. I choose countries whose markets are similarly important to the complainant as measured by the “export share” and include several from the respondent’s geographical region. I exclude countries engaged in similar WTO disputes.³⁴ Each WTO ruling has an implementation deadline which I use to split the sample into a “pre-treatment period” and a “post-treatment period,” lasting five years after the deadline. The goal is to create an accurate counterfactual—a synthetic control—from a weighted average of countries in the donor pool. The counterfactual is accurate when the export share for the synthetic control matches the respondent’s export share in the pre-treatment period.

The synthetic control is created with a two-part optimization process. First, each country in the donor pool receives a country-weight that optimizes the similarity between the respondent and the weighted average of the other countries on a number of covariates in the pre-treatment period (i.e. before the deadline). A country that

³³The synthetic control method is explained with mathematical detail in Appendix A.

³⁴See the online technical appendix.

looks similar to the respondent on these covariates receives more country-weight in the synthetic control. The country-weights are adjusted according to the relative importance of the covariates. Second, each covariate receives a covariate weight that minimizes the discrepancy in the pre-treatment period between the respondent's trade and the synthetic control trade, using the country-weights from the first optimization step. Covariates that are more important predictors of the respondent's trade receive more weight. Because this is a two-part optimization problem, a solution entails a set of country-weights and a set of covariate weights.

The covariates come from the World Development Indicators (WDI) database. I use gross domestic product (GDP), GDP per capita, annual GDP growth, agriculture value added, industry value added, manufacturing value added, services value added, trade dependence, and the unemployment rate.³⁵

The result is a synthetic control unit with trade patterns that closely resemble the respondent's actual trade in the pre-treatment period. Trade for the synthetic control is then projected into the following five years. This projection approximates the counterfactual—trade the respondent would have had in the absence of the WTO ruling. Using a difference-in-difference approach, I then compare the respondent's actual trade to the expected (synthetic control) trade.³⁶

The difference-in-difference approach allows me to generate a compliance score, denoted S , defined as the average yearly difference between the respondent's actual and expected trade in the post-treatment minus the average yearly difference in the pre-treatment period.³⁷ If this quantity is positive, the respondent's actual trade after

³⁵When calculating covariates for the EU, I average across all member countries, with membership updated by year in accordance with EU expansion.

³⁶Provided we assume the respondent's trade and the synthetic control trade follow parallel trends—are subjected to all the same systematic factors and shocks save the WTO ruling—this approach identifies the average treatment effect on the treated.

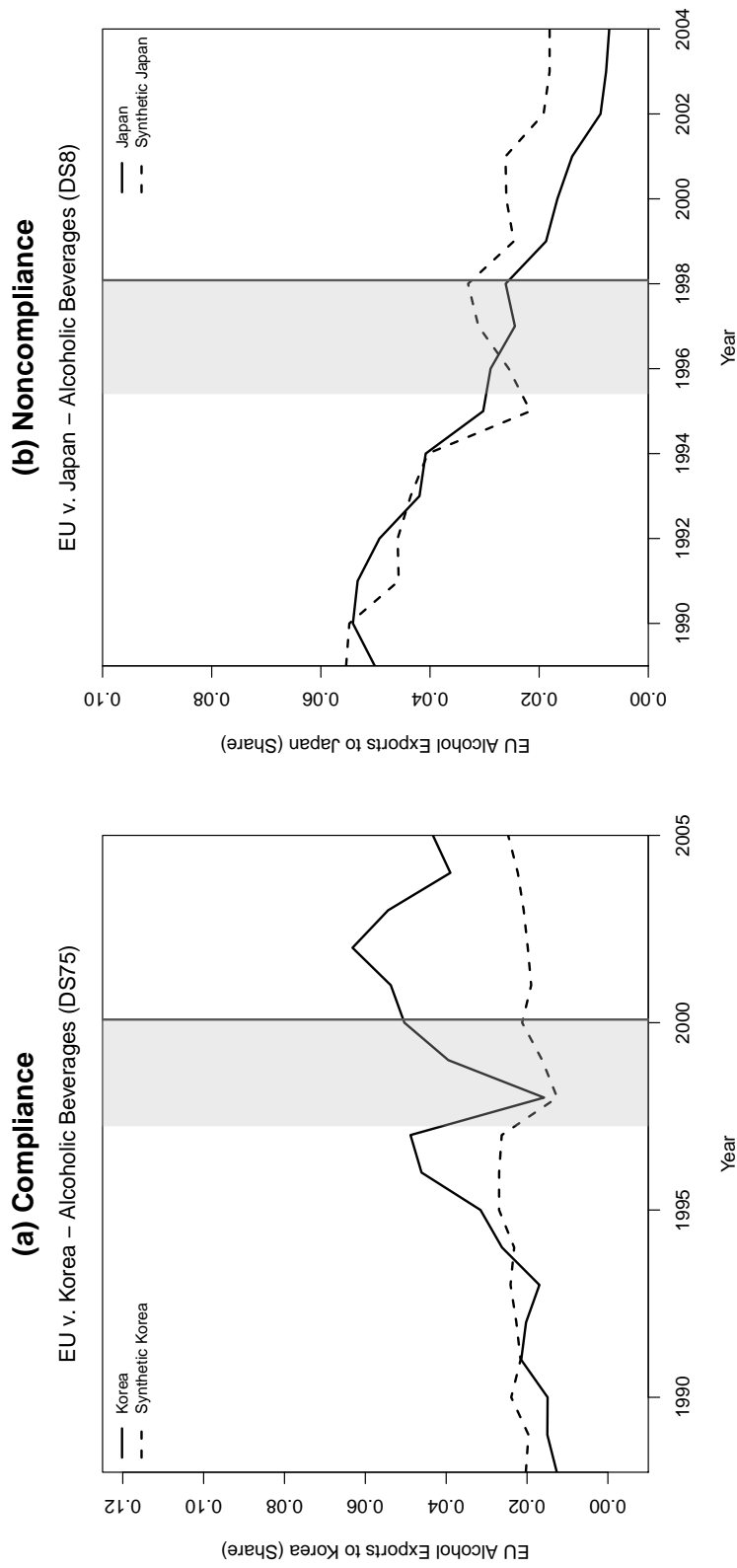
³⁷See the appendix.

the implementation deadline was higher than expected and I infer the respondent complied. Otherwise, I infer the respondent did not comply. I calculate a compliance score for each of the 125 WTO disputes and compute the standard deviation of these yearly measurements in the pre-treatment period, denoted d , to capture the stability of the estimator.

Compliance scores S are approximately normally distributed between -0.22 and 0.24 with a mean of -0.006 . On the upper bound, a compliance score of 0.24 indicates the ruling helped the complainant recover nearly one-quarter of its export market in the disputed product. Disputes vary widely in the dollar value at stake, but for the average dispute, this export share translates into roughly \$80 million in recovered trade. Nevertheless, the sample mean indicates that when WTO disputes are considered on average, the effect of adverse rulings on trade may be negligible.

The examples in Section 2.3.3 also illustrate my coding method. Figure 2.2 displays trade patterns for these WTO disputes. In the first case, the EU sued South Korea over its alcohol tax and won the lawsuit. The trends in European alcohol exports (product HS 2208) indicate that Korea complied. Figure 2.2(a) shows European exports to Korea rose relative to the synthetic control after the implementation deadline. SCM yields a positive, significant compliance score.

Figure 2.2: WTO Disputes with Compliance and Noncompliance Trade Patterns



Note: WTO disputes with compliance (a) and noncompliance (b). The vertical shaded areas cover the dispute duration from the request for consultations to the implementation deadline (bold vertical line). Synthetic control estimates were created with ‘Synth’ package in R.

In the second case, the EU sued Japan over a similar alcohol tax and won the lawsuit. It took many years for Japan to reform its tax system, long after the WTO's deadline for implementation had passed. The trade patterns indicate that Japan did not comply. Figure 2.2(b) demonstrates that European alcohol exports to Japan continued to fall, relative to the synthetic control. SCM yields a negative, significant compliance score.

2.4.4 Methodological Advantages

The synthetic control method has several advantages over standard matching or regression approaches. First, SCM creates more similar control units than standard matching techniques can achieve with a small, heterogeneous set of units. When a sample consists of few units, matching can be ineffective because treated units cannot be paired to control units without heroic extrapolations. The matching criteria can heavily influence the conclusions drawn (Smith and Todd 2005; Imai and Ratkovic 2014). By contrast, SCM is appropriate when the units of analysis are a few aggregate and heterogeneous entities like countries.

Second, SCM is transparent and flexible. It is transparent because it makes explicit the contribution of each comparison unit to the counterfactual. It is flexible because it allows each WTO dispute to have a separate covariate weighting that reflects the products and industries involved. As a result, SCM permits a direct analysis of similarities between the case of interest and the synthetic control. For example, in a dispute over computer chips, Korea's exports to the EU looked much like Korea's exports to Japan.³⁸ The covariates for GDP and industry value added received the most statistical weight in the calculation. In another example, a dispute over cigarettes, Honduras's exports to the Dominican Republic looked like a combi-

³⁸DS299: Korea v. EU - DRAMs.

nation of Honduras’s exports to Canada and to Costa Rica.³⁹ The most important covariates were agriculture value added and industry value added. These weightings comport with reasonable expectations about the sectors at stake in each dispute. Whereas conventional matching techniques impose the same covariate weights for all cases, SCM gives each case its own optimized covariate weighting.

Third, SCM helps to mitigate the problem of confounding factors. Many exogenous factors affect multiple countries in similar ways—for example, drought impacts an entire region. By choosing a qualitatively reasonable donor pool, I control for many such factors without explicitly using them as the basis for a match. This reduces the risk of omitted variable bias. Exogenous and unrelated changes are less likely to be driving my estimates.

Finally, SCM eliminates the possibility of extrapolation bias. Traditional regression analysis can lead to extrapolation outside the support of the data. By contrast, synthetic control units are computed as weighted averages (convex combinations) of other countries’ trade flows. SCM ensures all estimates are based on interpolation and well-supported by the data.

2.5 Analysis and Results

2.5.1 Dependent Variable - Compliance

The dependent variable, compliance, is generated from the synthetic control method, as described in Section 3.4. The compliance score S is the difference between actual and expected trade in the five years after the implementation deadline minus the difference in years before. While the sign of this quantity is reliable, the magnitude is sensitive to the number of years of trade data used to create the synthetic control unit, which in turn depends on data availability. This variability introduces noise into

³⁹DS302: Honduras v. Dominican Republic - Cigarettes.

the measurement. So I transform the compliance score S , and associated standard deviation d , according to two alternative coding rules.

Under the first coding rule, WTO disputes are differentiated by whether they have a negative or positive compliance score. Let COMPLIANCE_A be a binary variable:

$$\text{COMPLIANCE}_A = \begin{cases} 0 & \text{if } S \leq 0 \\ 1 & \text{if } S > 0 \end{cases}, \quad (2.1)$$

where 0 denotes noncompliance and 1 denotes compliance. With this coding rule, compliance occurs in 46% of the cases (58 of 125).

Under the second coding rule, I account for uncertainty associated with the compliance score. The standard deviation d reflects the overall variability in the match between the actual trade flows (treated unit) and the expected trade flows (synthetic control unit). Small standard deviations indicate the synthetic control unit precisely fits the observed data in the pre-treatment period and produces a more reliable compliance score. Cases are coded as compliance if the score is positive and larger than the standard deviation, noncompliance if the score is negative and larger in magnitude than the standard deviation, and inconclusive otherwise. Let COMPLIANCE_B be an ordinal variable:

$$\text{COMPLIANCE}_B = \begin{cases} 0 & \text{if } S \leq -d \\ 1 & \text{if } S \in (-d, d) \\ 2 & \text{if } S \geq d \end{cases}, \quad (2.2)$$

where 0 denotes noncompliance, 1 denotes an ambiguous outcome, and 2 denotes compliance. This metric incorporates the point estimate and the associated uncertainty. With this coding rule, I identify compliance in 35% of cases (44 of 125), noncompliance in 45.6% of cases (57 of 125) and ambiguous results in the remaining 19.2%.

2.5.2 Explanatory Variables and Controls

My explanatory variable is domestic veto players in the respondent government that arise from institutional checks and partisan opposition. I measure this as VETO POINTS using the Political Constraints Index (Henisz, 2002). It accounts for the number of independent branches of government, the extent of partisan alignment across branches of government, and preference heterogeneity within each legislative body. When measuring independent branches of government, it accounts for federal institutions. When measuring partisan alignment, it accounts for party composition and left/right preference which change over time. VETO POINTS range from zero (least constrained) to one (most constrained). This metric is ideal because it has comprehensive coverage across countries and over years in my sample and is widely-accepted among political scientists.⁴⁰ It allows me to draw comparisons between otherwise dissimilar countries, providing a unified summary of domestic obstacles. Where the European Union is the respondent, I use the average of the member countries' veto points, with membership updated by year.⁴¹

To control for international pressure, I focus on countries' revealed interest in the dispute. The number of third party countries serves as one proxy for international pressure. I code the number of THIRD PARTIES in each dispute using the Horn and Mavroidis (2008) data and WTO records. In my sample, 89 disputes have no third parties, 24 disputes have between one and three, and the remaining 18 have many third parties. The size of the complainant's economy is another manifestation of international pressure. Complainants with larger economies have a greater capacity

⁴⁰I use the Henisz index that excludes the judiciary. I also examine different measures from the Database of Political Institutions and control for democracy using the Polity IV Data set (Marshall and Jaggers, 2012). See the online technical appendix.

⁴¹In many of these WTO disputes, EU-wide policies are contested. Using the average of member countries' domestic veto points may understate the effective obstacles to EU compliance. The online technical appendix contains robustness checks.

to penalize respondents that defy a WTO ruling. I control for the COMPLAINANT GDP in the year the dispute was initiated.

As an additional control, I include the RESPONDENT GDP for the year the dispute was initiated. The GDP data come from the World Development Indicators database and these variables are normalized to improve comparability. I control for the extent of the % ADVERSE RULING, which measures the percentage of legal claims found in favor of the complainant. If the governments appeal the ruling, I account for the claims that were sustained on appeal. I use the Horn and Mavroidis (2008) dataset and WTO records, current to January 2013. Finally, I control for cases where the European Union is the respondent, EU RESPONDENT.

2.5.3 Results: Domestic Veto Points Hinder Compliance

2.5.3.1 Probit Model

When do WTO disputes lead to improvements in trade relations between the disputants? Across all model specifications, veto points are associated with a lower probability of compliance.⁴² Table 2.1 presents regression results with COMPLIANCE_A and a standard probit model. Veto points has a negative, statistically significant coefficient. The number of third party countries entering into the legal proceedings is positively associated with compliance, suggesting international pressure can factor into dispute outcomes. The extent of the adverse ruling is positively and significantly associated with compliance. It provides a useful proxy for the magnitude of the treatment. The more adverse, the more likely I am to detect trade patterns indicative of compliance.

The complainant and respondent countries' GDPs are not strongly associated with compliance. Once governments engage in litigation, their relative economic

⁴²See the online technical appendix for robustness tests with flexibility measures, alternative coding methods for compliance, additional control variables, groups of related disputes, etc.

power is not an informative predictor of compliance. This non-finding reinforces WTO advocates' claim that the legal process has an equalizing impact on the relations among countries, reducing the importance of power politics.⁴³

⁴³The EU does not have significantly different compliance rates.

Table 2.1: Probit Models for Compliance with WTO Rulings, 1995–2011

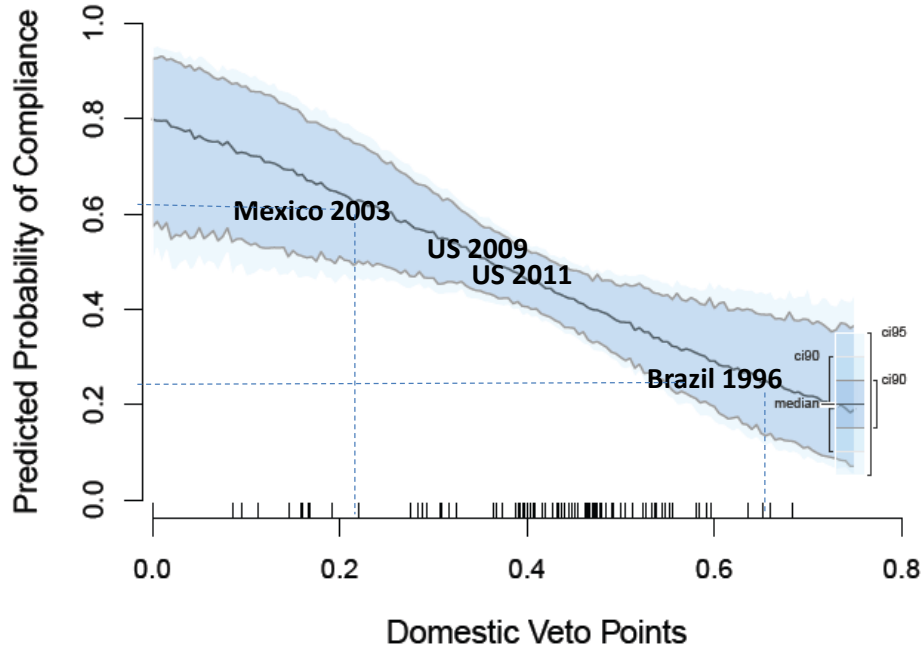
	COMPLIANCE _A						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Veto Points	-2.343** (1.042)	-2.203** (1.075)	-2.759** (1.083)	-2.279** (1.039)	-2.461** (1.049)	-2.100** (1.061)	-2.523** (1.180)
Third Parties		0.079 (0.054)					0.088 (0.055)
% Adverse Ruling			0.611* (0.336)				0.634* (0.360)
Respondent GDP				0.081 (0.104)			0.062 (0.125)
Complainant GDP					-0.081 (0.117)		-0.075 (0.125)
EU Respondent						-0.320 (0.313)	-0.239 (0.370)
Constant	0.870* (0.445)	0.735 (0.466)	0.623 (0.471)	0.819* (0.446)	0.919** (0.446)	0.825* (0.444)	0.447 (0.500)
Observations	125	125	125	125	125	125	125
Log Likelihood	-83.577	-82.473	-81.901	-83.274	-83.336	-83.048	-79.799

Notes: Compliance was coded using the synthetic control method with annual bilateral trade data for disputed products. Probit regressions are calculated with the ‘sampleSelection’ package in R. *p<0.1; **p<0.05; ***p<0.01.

One virtue of the probit model is its robustness to random coding errors. Random coding errors in the dependent variable will tend to attenuate (rather than inflate) the estimated effect of the explanatory variable (Cox and Snell, 1989). So random mis-assignment when I coded compliance should, in expectation, cause the regression to understate the magnitude of the coefficients.

Based on these estimates, Figure 2.3 shows the predicted effect of veto points on compliance. As veto points increase, the predicted probability of compliance decreases. Estimates use the probit model with all controls, holding variables at their means. The rug at the bottom of the plot shows the distribution of observations. Most disputes involve respondents with a moderate number of veto points.

Figure 2.3: Predicted Effect of Veto Points on Compliance



Note: Shading denotes 90% and 95% confidence intervals. Examples of veto points for selected country-years are plotted. The rug shows the distribution of veto points observations. Predictions are based on probit model with all controls. Estimates were created with ‘‘Synth’’ and ‘‘Zelig’’ packages in R.

Several examples are plotted. When Mexico was sued in 2003 over its imposition of anti-dumping duties, it had few veto points (0.284). Institutional constraints on the executive were modest. The Partido Acción Nacional (PAN) held the presidency and had pluralities in both houses of the legislature.⁴⁴ Mexico’s predicted probability of compliance was high (≈ 0.6) and in this instance it did comply. By contrast, when Brazil was sued in 1996 over its domestic aircraft program, it had many veto points (0.684). This reflects its federal system with many municipalities and the prevalence

⁴⁴In 2003 the PAN held 38.11% of the seats in the senate and 38.24% of the seats in the lower house compared to smaller shares held by the main opposition party.

of coalition government among multiple political parties.⁴⁵ Brazil’s predicted probability of compliance was low (≈ 0.2) and it did not comply. Most of the variation in veto points is across countries, which have many fixed institutions. But there is also some variation across time for individual countries, due to partisan shifts. For example, in 2009 the United States has fewer veto points (0.397) than it did in 2011 (0.414) when the Republican party gained a majority of seats in the House of Representatives while the Democratic party maintained control of the Senate and Presidency.

2.5.3.2 Ordered Probit Model

Domestic political divisions are strongly associated with less compliance, even using the alternative coding rule. Here, COMPLIANCE_B is an ordinal variable that incorporates the uncertainty associated with SCM compliance scores and I fit an ordered multinomial probit model. Table 2.2 displays results.

⁴⁵The Brazilian Constitution treats its 5,570 municipalities as parts of the Federation, each with autonomous local government, and not simply dependent subdivisions of its 26 states and Federal District.

Table 2.2: Ordered Probit Models and Linear Model of Compliance with WTO Rulings, 1995–2011

	COMPLIANCE _B				COMPLIANCE SCORE (S)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Veto Points	−1.905* (0.973)	−2.383** (0.977)	−2.033** (0.947)	−2.281** (0.962)	−1.797* (0.964)	−2.077* (1.072)	−0.040 (0.057)
Third Parties	0.085* (0.051)					0.099* (0.052)	0.003 (0.003)
% Adverse Ruling		0.450 (0.305)				0.409 (0.328)	0.010 (0.019)
Respondent GDP			0.060 (0.096)			−0.009 (0.114)	−0.004 (0.007)
Complainant GDP				−0.115 (0.107)		−0.119 (0.116)	−0.006 (0.006)
EU Respondent					−0.395 (0.292)	−0.428 (0.342)	−0.023 (0.019)
Constant (0:1)	−0.971 (0.409)	−0.821 (0.427)	−0.789 (0.430)	−0.933 (0.412)	−1.055 (0.414)	−0.922 (0.408)	
Constant (1:2)	−0.468 (0.405)	−0.309 (0.425)	−0.427 (0.428)	−0.428 (0.408)	−0.548 (0.409)	−0.413 (0.405)	
Constant							0.007 (0.025)
Compliance Measure	(0/1/2)	(0/1/2)	(0/1/2)	(0/1/2)	(0/1/2)	(0/1/2)	Continuous
Model	OP	OP	OP	OP	OP	OP	OLS
Observations	125	125	125	125	125	125	125
Log Likelihood	−126.353	−126.679	−127.580	−127.201	−126.847	−123.578	

Notes: Compliance was coded using the synthetic control method with annual bilateral trade data for disputed products. Ordered probit regressions are calculated with the "MASS" package in R. *p<0.1; **p<0.05; ***p<0.01.

Table 2.2 lends additional support to the hypothesis and reinforces the findings above. A respondent government with more veto points is less likely to comply with an adverse WTO ruling. Using the compliance scores directly in the regression (column 7, OLS) confirms there is a tendency, if statistically insignificant, for countries with more veto points to comply less.

Predicted effects demonstrate the magnitude of the findings. If an average respondent government in a WTO dispute were to increase its domestic veto points from the least to the most, its predicted probability of compliance *decreases* by 0.44. The estimate is based on the ordered probit model with all control variables (column 6) and conditions on a transition from the ambiguous outcome to compliance.

2.5.4 Robustness

2.5.4.1 Reverse Causality

Endogeneity bias might arise if international trade disputes affect domestic politics. Scholars have shown trade affects domestic political cleavages (Rogowski, 1989) and that international conditions can force a country to adapt its trade policy process (e.g. US fast track negotiating authority). This form of reverse causality is not a serious concern in my analysis. Trade disputes occur over a relatively short time-frame: usually less than five years. Changes in trade policy have a rapid impact on trade flows. By contrast, domestic political divisions, as measured by veto points, change slowly over time. They reflect (1) the institutional separation of power which may change only a few times during a country's modern lifespan and (2) partisan divisions which are determined by unrelated and powerful macro-political issues. WTO disputes are unlikely to affect the respondent's veto points.

2.5.4.2 Selection Effects from Early Settlement

WTO disputes with adverse rulings are only a subset of the legal cases brought to the Dispute Settlement Mechanism. More than half of the disputes are resolved without any legal verdict because the disputants settle early. This reflects the WTO's aim to resolve trade disputes during consultations, before governments resort to costly litigation and compliance becomes an issue. Selection into litigation is not random, and potentially correlated with the outcome—compliance—raising the possibility of selection bias.⁴⁶

The direction of potential selection bias is *a priori* unclear. On the one hand, disputes that are especially intractable may be more likely to require a panel ruling and less likely to result in compliance.^{47,48} On the other hand, lawsuits against particularly obstinate respondents may be resolved prior to litigation if the complainant, realizing litigation is futile, capitulates. This type of selection effect is notoriously difficult to overcome using observational data.⁴⁹ The outcomes I observe may either understate or overstate the theoretical compliance rate, had all disputes gone through litigation.

As a first step, I examine the selection process with a series of Heckman selection models (Heckman, 1979, 1990), shown in the Appendix. My results are robust in that I do not find any evidence that contradicts my theory: domestic veto points are still strongly associated with noncompliance. In these models, the selection stage

⁴⁶Because nearly every WTO ruling is adverse, the important comparison group is the set of disputes that were settled early, without a legal verdict at all.

⁴⁷This expectation is probabilistic because potential complainants are not always well-informed and the respondent's policy preferences change with exogenous political and economic shocks. Even respondents with many veto players have a small chance of complying.

⁴⁸In addition, countries may choose whether to violate GATT/WTO rules based on expectations about retaliation (Bown, 2004a).

⁴⁹As strategic actors, governments in an international trade dispute have every incentive to anticipate their opponent's behavior and adjust their actions.

estimates the effect of covariates on the probability the WTO issues an adverse ruling and the outcome stage estimates the probability of compliance, conditional on the adverse ruling. I use a number of different variables for identification. However, this robustness test is not decisive because the identification is imperfect.

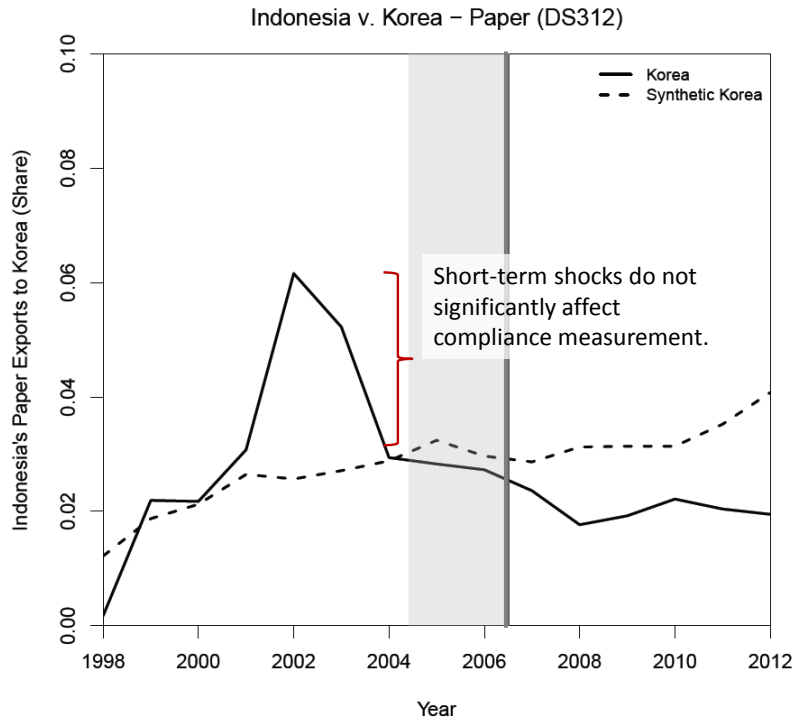
2.5.4.3 Endogenous Dispute Timing

Governments prefer to initiate trade disputes when conditions are favorable. They may wait to file their complaints about treaty violations until there are exogenous shocks that lead to declines in trade. The WTO usually requires the complainant to demonstrate harm to its domestic industry and declining trade provides such evidence. If complainants prefer to file their disputes when there are exogenous declines in trade, dispute initiation may be endogenous to trade fluctuations.

Dispute timing could bias my estimates. If complaints are filed when there are negative shocks, once the exogenous shock passes trade should increase, even absent a WTO ruling. Then I might mistakenly attribute to the WTO an increase in trade that is due to the unrelated passing shock. In other words, endogenous timing would bias my estimates toward compliance and thereby overestimate the efficacy of the WTO.

My methodology is robust to dispute timing because the synthetic control unit is constructed from many years of trade patterns for several countries and short-term fluctuations are moderated. Countries experience temporary import surges that pass. When the surge is viewed over a short time horizon, it might appear as though trade declined. However, viewed over a longer time span, it is clear the decline merely restores pre-surge trade levels. My methodology is relatively impervious to the stochastic fluctuations that complainants exploit when initiating WTO disputes. I am still able to detect instances of noncompliance.

Figure 2.4: Robustness to Dispute Timing



Note: The vertical shaded area covers the dispute duration from the request for consultations to the implementation deadline (vertical line). Synthetic control estimates were created with ‘‘Synth’’ package in R.

For example, Figure 2.4 shows trade flows in a dispute between Indonesia and Korea⁵⁰. Korea experienced an import surge around 2002 and imposed a trade barrier that broke WTO rules. Indonesian exports to Korea dropped precipitously and Indonesia filed a formal complaint at the WTO. Viewed over the two years leading up to the dispute, it would appear that Indonesian exports suffered a huge loss from Korea’s policy because the exogenous shock magnified the apparent impact of the trade barrier. Yet viewed over a longer time span, it is clear that Indonesian exports to Korea were unusually high before the dispute and Korea’s barrier had only a modest impact. My SCM estimates are robust because I account for the

⁵⁰DS312: Indonesia v. Korea–Paper

longer time horizon and use data from many countries to construct the control. In Figure 2.4, the control unit trade (dashed line) does not increase during the import surge. As a result, I am still able to show that Korea did not comply with the WTO ruling. This matches the legal record—compliance proceedings showed Korea failed to implement the ruling. Dispute timing does not appear to confound my estimation. The synthetic control method is a reliable way to measure compliance.

2.6 Conclusion

While the World Trade Organization’s Dispute Settlement Mechanism is central to the multilateral trade regime, relatively little is known about its effect on international cooperation. This paper evaluates the impact of adverse WTO rulings on trade flows between disputing countries by applying the method of synthetic case control. In every dispute, I estimate the causal effect of an adverse WTO ruling by constructing an estimated counterfactual against which to gauge trade fluctuations. The counterfactual incorporates specific product-level dyadic trade data for several years leading up to and following the ruling. Each is created with an optimal combination of economic control variables that reflects the products and issues in that particular case. Increases in actual trade relative to expected trade indicates the respondent government complied with the ruling. Using this methodology, I measure compliance for all 125 WTO disputes with adverse rulings between 1995 and 2011.

When its record is evaluated in the aggregate, it might appear that the WTO has had little success in restoring trade between disputing countries. However, my results demonstrate that many disputes actually prompt significant increases in trade, relative to expected levels. Other disputes have no discernible effect. To my knowledge, this is the first study to estimate the impact of WTO rulings and to account for the substantial heterogeneity in outcomes. My results show domestic politics explain a respondent government’s behavior when it faces an adverse ruling. Conditional on

violating a treaty, the more domestic veto points a government has, the less likely it is to return to compliance. This pattern is significant and persists even when the respondent faces international pressure to comply. It is robust across model specifications and coding schemes.

My findings have implications for research on the optimal design of international institutions. Many scholars have highlighted the advantages of flexibility mechanisms which provide a safety valve to countries facing acute domestic pressure to defy their international commitments. Governments are less likely to abandon an institution altogether when they are permitted temporary transgressions (Rosendorff and Milner, 2001; Carrubba, 2005; Rosendorff, 2005; Johns, 2014). The models that link flexibility to stability typically treat states as unitary actors that can violate rules and then easily return to their international commitments. However, opening the black box of domestic politics reveals that returning to compliance may not be so simple; multiple veto players can obstruct the process. This implies a potential hazard of institutional flexibility: it only promotes long-term cooperation when countries return to compliance after the temporary pressure to violate passes. When veto players lock in treaty violations, flexibility mechanisms may fail to restore cooperation in the long-term, undermining the institution's stability. Dispute settlement mechanisms may be particularly vulnerable due to the publicizing nature of litigation which forces the government into an often-fraught domestic political process. At this point, treaty violations become difficult to reverse.

Political scientists have long known that divisions in a government can obstruct policy change. The impact of institutional checks and balances and partisan conflict—veto players—also reaches beyond national borders. Veto players have been shown to impact the types of international commitments countries make and keep (Milner and Rosendorff, 1996, 1997; Rickard, 2010; Mansfield and Milner, 2012). Institutional checks and partisan opposition are integral to democracy. Many studies therefore conclude that democracies, with their multiple veto players and voter-based audience

costs, comply more with international commitments (“first-order compliance”). In addition, democracies are thought to allow for the creation and growth of interest groups that support international cooperation (e.g. Finnemore and Sikkink 1998; Dai 2005; Simmons 2009).

That key features of democratic politics sometimes obstruct compliance sits uneasily with the prevailing view in the international relations literature. My findings point to a more nuanced relationship, suggesting that democratic politics can have cross-cutting effects on international cooperation. Veto players can actually decrease a government’s likelihood of complying with international legal rulings (“second-order compliance”). With domestic institutional divisions and partisan opposition blocking policy change, even trade policies that violate international obligations can be locked into place. Dispute settlement mechanisms are often unable to compel governments to reverse these violations. Ultimately, international dispute settlement may only be as effective as domestic politics allows.

CHAPTER 3

The European Court of Justice and EU Economic Cooperation: Do Domestic Political Divisions Hinder Reform?

3.1 Introduction

Do legal rulings from the European Court of Justice promote European economic cooperation? A long history of European Union scholarship suggests the judicial system has played an instrumental role, leading to the integration of Europe. Yet the evidence is still sparse with respect to the most fundamental dimension of European cooperation: trade. In this chapter, I argue that the European Court of Justice (ECJ) has promoted further economic integration but that its impact remains highly constrained by domestic politics in the member states. When some member countries of the European Union lose lawsuits on trade-related issues, they subsequently open their economies to more intra-European trade. Others show no such evidence of change. Some of the variation is explained by the domestic political constraints European governments face. When the country has divisive domestic politics—as measured by domestic veto points—it is less likely to show trade increases in the wake of adverse legal rulings than when its domestic politics are more unified. These findings suggest that when the European Court of Justice orders countries to follow through on their commitments to European economic integration, domestic political divisions inhibit reform.

This chapter uses a combination of Bayesian and classical (frequentist) statistics to analyze patterns of European economic cooperation in the wake of European Court of Justice (ECJ) rulings. When the adverse ECJ rulings are *interacted* with domestic veto points, they are negatively associated with intra-EU trade. The correlation is consistent when I examine different temporal lags, control variables, and model specifications. This study lends insight into the role of the European Court of Justice in promoting European economic integration.

A major challenge to studying international institutions is distinguishing between institutional effectiveness and selection effects. Because countries voluntarily enter these institutions, it is difficult to identify whether countries comply with their obligations because the institution compels them to do so, or because that is what they would have done in the institution's absence anyway. Infringement disputes at the European Court of Justice arise when a national government violates its obligations under EU law, despite knowing its responsibilities and despite pressure from the Commission to implement EC directives. By virtue of its resisting EU law, retaining an illegal policy despite the threat of lawsuit, and then losing the ECJ ruling, that national government makes clear its preference to not comply. If the government nevertheless implements the ECJ ruling through appropriate domestic policy, it complies, despite these preferences. When a government complies with a ruling by correcting an initial violation, we may infer that the ECJ can indeed promote European cooperation.

As the European Union has deepened its membership obligations, domestic governments have been required to implement numerous EU directives, legislative acts that members enact through domestic law help achieve the goals of the EU. Frequently member states resist, prompting the central administrative body, the European Commission to sue. In the majority of these lawsuits, the court rules against the member state and orders reform. Yet because European economic integration is often a politically volatile issue, domestic governments have confronted substan-

tial obstacles to implementing such rulings. The more political constraints within a government from both institutional checks and balances and from partisan discord, the more significant the obstacles and the less likely the domestic government is to respond to an adverse ruling and open its economy to further EU commerce. Conditional on initial violations, European members that face constrained domestic politics are less likely to follow through on ECJ rulings in substantial ways that increase EU integration. The findings suggest that domestic political divisions *hinder* economic cooperation in the European Union (EU).

The remainder of the chapter proceeds as follows. The second section discusses the role of the European Court of Justice in promoting European economic integration and argues that domestic politics creates obstacles to reform. The third section presents data and statistical models to evaluate the theory. The remaining sections present results, discuss the findings, and conclude.

3.2 Regulation of Trade, ECJ Rulings, and Domestic Obstacles

3.2.1 Regulation of Trade

Originally a free-trade agreement founded in 1957 among European countries, the European Union¹ is at its heart an economic organization aimed at promoting commerce between member countries. The member states of the European Union (EU) are obligated to open their economies to one another. By the European Union's own description, "the [founding] Treaty provided for the establishment of a common market, a customs union and common policies....the Community's primary mission is to

¹The organization known today as the European Union was created by the Treaty of Rome as the European Economic Community or "Common Market" among Belgium, France, Germany, Italy, Luxembourg and the Netherlands. It has gone through several expansions and to date includes 28 members.

create a common market” and provide institutions to achieve this objective.² Over the past four decades, the EU has evolved into a common market with economic policies that ensure the free movement of most goods, services, labor and capital within the territory. EU obligations further include a common currency for many member states and uniform policies on trade with external economies.³

While the European single market is founded on a series of treaties to which each member country commits, it is further enabled by legal orders—“directives”—which the European Commission, the central administrative branch of the EU, periodically issues. Many directives aim to eliminate barriers to European commerce in specific areas of EU integration, for example by standardizing tax rules or competition policy across member states. Others address policies that reduce unfair competition within certain industries or sectors, for example directives that eliminate the use of certain pesticides in agriculture or require environmental standards in waste disposal. Directives ordinarily apply to all members of the EU and must be implemented at the national level by member states themselves through domestic law or policy.⁴ It is the responsibility of the member governments to report to the Commission how they implemented the directive or how their laws already meet the requirements.

While European countries all commit to adopt EU directives on economic integration, there is variation in how well they implement these orders. EU directives on topics like the internal market, taxation, the customs union, and competition are often the subjects of controversy. Sometimes, national governments staunchly resist the court, especially when there are influential domestic groups that oppose implementing the directive. For example, the controversial investment services directive

²“Summaries of EU Legislation,” *Europa*, http://europa.eu/legislation_summaries/institutional_affairs/treaties/treaties_eec_en.htm.

³To date, eighteen out of 28 EU members uses the common currency, the Euro.

⁴“The EU Single Market: Policy Framework,” *Europa*, http://ec.europa.eu/internal_market.

in 1993 brought significant new security regulations that many firms opposed. Other times, the obstacles to implementation are institutional if, for example, countries lack bureaucratic capacity to make the necessary changes (Börzel et al., 2010). So although member states are legally obligated to implement EU directives correctly and in a timely manner, they sometimes fail to do so. In such instances, the European Commission can investigate and prosecute the matter. Continued failure to implement the directive often provokes a lawsuit at the European Court of Justice, the main judicial body of the EU. This is an “infringement dispute,” an instance where European bureaucracy sues a member state for not complying with its primary commitments.

3.2.2 ECJ Infringement Disputes

As the European Union expands and deepens its economic obligations, it has kept the European Court of Justice busy. The ECJ has adjudicated nearly 1,700 infringement disputes between 1978 and 1999 and, according to many scholars, has been instrumental in the process of European integration (Sweet and Sandholtz, 1997; Garrett, Kelemen, and Schulz, 1998; Sweet and Caporaso, 1998; Tallberg, 2000; Panke, 2007). In the vast majority of these cases—88% of disputes over trade-related issues—the ECJ rules in favor of the European Commission and thus against the member state. These rulings almost always mandate greater trade liberalization—and less autonomy for the national government in setting its own policies that might otherwise restrict intra-EU commerce. Between 1988 and 1999, the years for which comprehensive data are available, there were 225 adverse trade-related ECJ rulings against the fourteen core EU countries.⁵ These are the instances in which the ECJ decided the member was guilty of failing to implement a trade-related directive.

⁵The EU member states in this analysis are: Austria*, Belgium, Denmark, Finland*, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden*, and the United Kingdom. Three of the countries (denoted *) joined the EU in 1995.

After the ECJ issues an adverse ruling, the defendant country is required to implement it through domestic policy. The country must follow through on its primary obligation to follow European directives, despite its initial reluctance that provoked the dispute in the first place. On first inspection, it might appear that the ECJ's compliance record is strong: few countries are prosecuted a second time for failing to implement the ECJ's verdict. Between 1988 and 1999, for example, only 13% of the ECJ cases with adverse rulings were re-investigated for continued violations. Yet a closer look at the ECJ dispute process suggests this apparent legal *de jure* compliance rate is not a reliable indicator of *de facto* compliance, the meaningful changes in behavior that member countries are obligated to make.

There are many reasons that the *de jure* compliance rate is likely to overstate the actual *de facto* compliance rate. First, the Commission may choose to drop a lawsuit even though a defendant government did not comply. When the defendant is particularly obstinate, the Commission may decide the lawsuit is futile and choose not to bear the cost of continued litigation or the risk of a damaged reputation: persistent non-compliance might undermine the legitimacy of the Commission and possibly the European Court of Justice in the eyes of the other member states. Second, the defendant government may implement a new law or policy that *appears* compliant but does little to accomplish the economic outcome represented by the directive, especially if the defendant government does not enforce its new law.⁶ For example, a defendant government could legally comply by adopting more stringent standards for investment services but then fail to penalize firms that break the rules. When compliance is costly, countries may find ways to circumvent the legal ruling or ignore it altogether.

I focus on the *economic impact* of ECJ infringement rulings on intra-EU commerce. ECJ rulings that favor the Commission promote European economic integra-

⁶Moreover, Garrett, Kelemen, and Schulz (1998) and Carrubba, Gabel, and Hankla (2008) argue that the ECJ is more apt to issue adverse rulings in cases where it expects compliance to be likely.

tion by requiring members to adopt policies that facilitate commerce across national borders. If rulings are faithfully implemented—if members comply—we should expect to see trade increase. By complying with a trade-liberalizing ruling, the member becomes more economically integrated into the EU and this should be exhibited by larger portion of its imports coming from other EU countries. Whereas legal compliance and policy changes are difficult to measure across countries and with respect to the many trade-related topics, trade flows present a relatively objective metric. When adverse trade-related ECJ rulings lead the defendant to import more from other EU countries in subsequent years, all else equal, I infer that the legal process is promoting European economic cooperation. Conversely, a negative correlation between adverse trade-related rulings and subsequent intra-EU trade suggests that the legal process is not promoting European economic integration.

More than most international courts, the ECJ has some ability to enforce its rulings. Through two EU doctrines, “direct effect” and “supremacy,” ECJ rulings are directly placed within the national legal orders and any national laws that conflict with them are meant to be removed at once. This makes ECJ rulings binding on member states and automatically applicable. When a country defies the court by failing to comply, the ECJ can impose penalties which are retroactive and apply for the full duration of the violation. In the rare cases where the defendant continues to defy, the European Commission can call on the ECJ to impose additional financial sanctions. These sanctions are meant to be large enough so as to have a deterrent effect.⁷ Nevertheless, as an international court the ECJ’s ability to enforce is limited to financial sanctions and even the steps required to impose these penalties are many. It cannot directly force a member state to comply.⁸

⁷Article 260 of the *Treaty on the Functioning of the European Union* and Article 106a of the *Euratom Treaty*.

⁸For example, when the ECJ struck down France’s ban on British beef, the French government defied the Court. One British newspaper reported: “The European Commission admitted it was powerless to take swift action against France over the continued blockage of British beef...it would

A state's response to adverse ECJ rulings on infringement cases offers a rare view into institutional effectiveness. A major obstacle to studying the effectiveness of international institutions is that it is difficult to separate causal impact from selection effects. Presumably, countries choose to join international institutions that align with their interests. If they comply with those commitments, first-order compliance, they may simply be adopting behavior they would have chosen anyway, in the absence of the institution (Downs, Rocke, and Barsoom, 1996; Von Stein, 2005). Second-order compliance, on the other hand, can be particularly revealing. These are instances where national governments demonstrate an explicit preference to violate their primary obligations—we know this because they failed to implement the EC directive and lost an ECJ ruling. When, despite this preference, the government complies with the ECJ ruling, we can infer the institution had some influence. In this respect, compliance with ECJ rulings in infringement disputes lends insight into the effectiveness of ECJ in promoting European economic cooperation.

Note that infringement disputes, the focus of this analysis, are distinct from another major category of ECJ lawsuits, preliminary references. Through the preliminary reference system, lawsuits filed under the member states' national court system are referred to the ECJ when they bring to bear issues of European law. The ECJ rulings are then incorporated and applied through the member state's domestic legal system. I do not evaluate preliminary references in this study for two reasons. First, they have wide variation in plaintiffs which can include individual citizens, corporations, industry groups, etc. so it is difficult to consistently map rulings onto trade liberalizing policies or compare across lawsuits. Infringement disputes, by contrast, all have the same plaintiff, the Commission. Second, the implementation process for preliminary references is internalized through each country's own judicial

take years to go back to court to seek fines and compensation.” And the French government's delay may have been driven by domestic political factors. Agriculture Minister Lord Whitty said: “The French are clearly out of order. However [they] may not move until after the French elections.” See: “No action against France for months in beef row” *The Journal, Newcastle UK*, March 14, 2002.

system, blurring the line between European level rulings and domestic policy. As such, preliminary references may be the “easy” cases in which to show an impact of ECJ rulings on economic integration while infringement disputes are the “difficult” cases.

3.2.3 Domestic Constraints

Domestic politics can interfere with second-order compliance by making it more difficult for governments to implement rulings from the European Court of Justice. Domestic groups may oppose the policy changes that the ECJ rulings require and then pressure their government to defy the Court. For example, when the ECJ issued a ruling about gender equality in insurance and pensions,⁹ insurance companies throughout the EU opposed compliance. One major UK insurance company stated that “the ECJ had made a horrible mistake” and warned “there would be huge costs to the industry repricing and updating systems...everyone loses.”¹⁰ Another said “there was a possibility the [UK] government could resist changing the corresponding legislation” by ignoring the ECJ judgment and dragging the process out for years.¹¹ With an estimated £900 million in costs to the UK industry alone, the government was under pressure to defy the ruling.¹²

Political constraints within the defendant government can obstruct policy change, particularly policies that shape international cooperation. Indeed, domestic political constraints have been linked to commercial openness (Henisz and Mansfield, 2006),

⁹This case was a preliminary reference from Belgium’s domestic courts, C-236/09, *Test-Achats v. Council of Ministers* but nevertheless obliged other EU members to reform their insurance and pension policies.

¹⁰“European sex equality ruling will lead to insurance chaos,” *The Scotsman: Business Edition*, March 2, 2011.

¹¹*Ibid.*

¹²“Madness of EU pension ruling,” *Scottish Express*, February 28, 2011.

trade agreements (Mansfield and Milner, 2012) and patterns of infringements in the European Union (Mbaye, 2001). Governments with substantial checks and balances have high institutional obstacles for policy reform. And governments with divisive partisan politics, e.g. strong opposition parties, can lead to discord that inhibits policy changes. Institutional checks and balances and partisan opposition create “veto points” in domestic government. These veto points make it more difficult for governments to implement ECJ rulings because they introduce more opportunities for domestic groups that oppose compliance to obstruct the process.

Domestic politics is often deeply divided over opinions on European Union integration and the government’s response to ECJ rulings. Concerning a high-profile dispute over labor restrictions, one British newspaper reported: “Britain has appealed at the European Court but is expected to lose. Close allies of the prime minister are advising him to ignore the judgment if that happens, and challenge Labour over whether it would back him or ‘cave in’ to Brussels”¹³ The more veto points in domestic government, the more likely the implementation process is to be stalled by political opposition and the less likely the government is to comply.

All else equal, a defendant country that complies with adverse ECJ rulings should open its borders to more European integration, leading to an increase in its intra-EU imports. But the more domestic veto points the domestic government has, the less responsive it should be to these adverse ECJ rulings. Thus the *interaction* of domestic veto points and trade-liberalizing ECJ rulings should be associated with less EU integration. My approach combines insights from previous literature that examines the degree to which ECJ rulings can liberalize trade (Gabel et al., 2012) and the impact of domestic political institutions on international trade cooperation (e.g. Milner and Rosendorff (1997); Rosendorff and Milner (2001); Goldstein and Martin (2000); Rickard (2010, 2012))

¹³ “Ministers prepare to defy Euro-Laws” *The Sunday Times*, London, October 15, 1995.

3.3 Data and Model

3.3.1 Dependent Variable: Intra-EU Trade

I evaluate economic integration using trade patterns *within* the European Union. The unit of analysis is the country-year, with every EU member observed annually over the entire period 1988 to 2010. I measure economic integration as the IMPORT SHARE, each member's imports from other EU countries in a given year divided by its total imports from the world that year. The import share is a useful way to measure European economic integration for several reasons. First, it captures each country's relative preference for intra-EU commerce. I am primarily interested in European countries' compliance with legal commitments for economic integration. Therefore, I aim to measure countries' European economic integration relative to their overall integration in the global economy. Second, it controls for system-wide economic shocks that uniformly affect trade volumes (like the 2009 financial crisis). This stands in contrast to previous ECJ studies, which have examined the effects of rulings on total intra-EU trade. I focus on *imports* because these are the trade flows that are directly addressed by the ECJ's pro-liberalizing rulings—cases where a defendant government must reduce barriers to imports. By contrast, these rulings do not have clear implications for exports.

There should be a multi-year lag between when the ECJ issues a ruling on when one might expect to observe an impact on the defendant country's imports. Some scholars have found an approximately two-year delay between the referral of a dispute to the ECJ and the ruling while others have indicated that even longer delays are probable before any impact on imports should be discernible (Alter, 2001; Gabel et al., 2012). Accordingly, I test lags of 2, 3, and 4 years.

Let $j \in J = \{1, 2, \dots, 15\}$ index the European Union countries and $t \in \{1988, 1989, \dots, 2010\}$ index years. Let y_{jkt} denote imports into country j from country k in year t and let

Z_{jt} denote total imports into country j from the entire world in year t . Then the dependent variable Y_{jt} , import share, is calculated as:

$$Y_{jt} = \frac{1}{Z_{jt}} \sum_{k \neq j \in J, t} y_{jkt} \quad (3.1)$$

Trade data are from EuroStat (2014) and include all goods.¹⁴

One possible disadvantage of this metric is that it will conflate economic specialization with economic integration. If a country has an economic profile that particularly *differentiates* it from the rest of Europe relative to the rest of the world, it will have a larger import share. One cannot tease apart which part of the import share is due to European economic integration and which part is due to market differentiation. However, because market differentiation changes little over time, changes in intra-EU trade shares are more likely attributable to changes in economic integration. By using a lagged dependent variable in the models below, I aim to control for this possible problem.¹⁵

Figure 3.1 shows the trends in imports for each country, from 1988 to 2010. This is the value of imports from the European Union, measured in Euros (a) and as a share of total imports from the world (b). The data span two structural changes to the European Union. The 1995 enlargement brought Austria, Finland and Sweden into the EU and the 1999 adoption of the EuroZone brought a common currency.

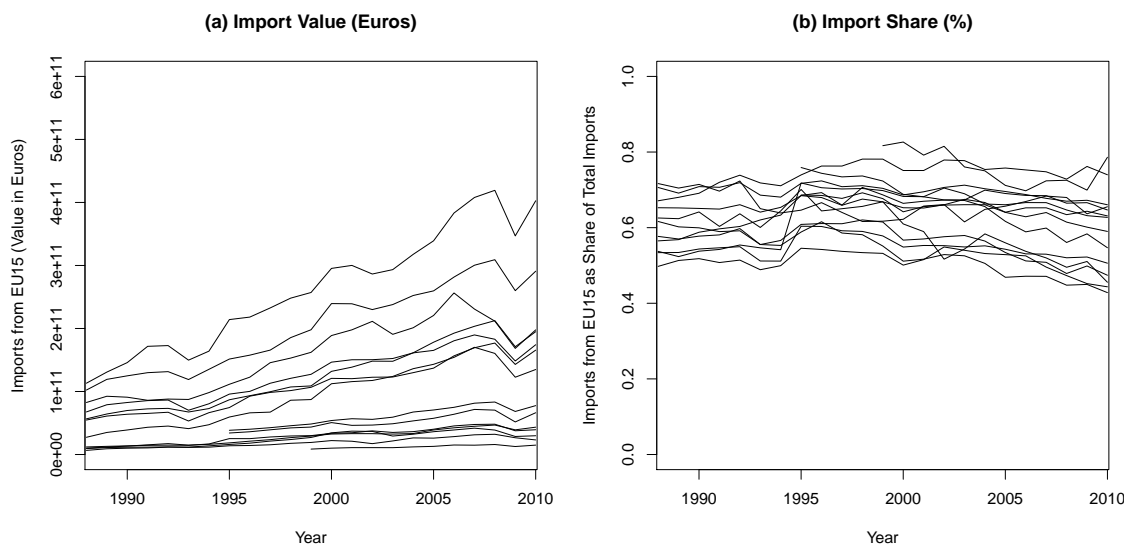
The value of each member's imports from the EU (measured in €) has tended to increase over time, alongside overall economic growth trends. Between 1988 and 2010, the notable exception occurs during the 2008 financial crisis, when there was a substantial drop in trade for nearly all European countries. Import shares (measured

¹⁴I do not account for trade in services because of data availability issues.

¹⁵In various model specifications, I use country-random intercepts that should further ameliorate the problem.

as %) have remained fairly steady over time. Some countries have become more integrated in to the European Union but most display little temporal variation.

Figure 3.1: Intra-EU Imports Over Time by Country



Note: Data from EuroStat. Import share is a country's imports from EU15 divided by total imports from world. From 1988 to 1999, EuroStat aggregated Belgium and Luxembourg, reporting trade values for a single entity "BelLux." After 1999, Belgium and Luxembourg are reported separately.

3.3.2 Explanatory Variables: ECJ Rulings and Domestic Veto Players

The first explanatory variable is each country's number of adverse ECJ rulings on a trade-related infringement dispute, by year. All country-years are included, even those in which a member did not experience an adverse ruling. I restricted the analysis to only infringement disputes that concerned trade and intra-EU commerce and where the Court ruled *against* the member state. This includes disputes over policies relating to: internal market, budget, taxation and customs union, competition policy, agriculture, enterprise and industry, and economic or financial affairs. In most years, many EU members have no instances of losing an ECJ infringement lawsuit

over trade-related directives. Unsurprisingly, Austria, Luxembourg, and Sweden, have the fewest incidents. On the other hand, Germany, Greece, and Italy experienced the highest number of adverse ECJ rulings. In a single year, France lost 11 infringement disputes over trade issues and in the typical year, Italy loses 3 or 4 such lawsuits. These data are summarized in Table 3.1. All information about the infringement disputes, legal verdicts, and subsequent noncompliance, if the country was investigated for failure to implement the ruling, come from the Database on Infringement Incidents in the European Union, coded from the European Commission's own comprehensive records (Börzel and Knoll, 2012; Börzel, 2001; Börzel, Hofmann, and Panke, 2012). In the regression tables below, the number of adverse trade-related rulings are denoted by ATR_{t-n-1} and are lagged to allow for delays between the ruling and implementation.

Table 3.1: Frequency of Adverse ECJ Rulings per Year By Country

	ADVERSE RULINGS PER YEAR									TOTAL
	None	1	2	3	4	5	6	7	11	
<i>Country</i>	<i>Frequency - Number of Years</i>									
AUSTRIA	13	2	0	0	0	0	0	0	0	2
BELGIUM	11	2	4	2	2	1	0	1	0	36
GERMANY	8	9	5	1	0	0	0	0	0	22
DENMARK	19	2	2	0	0	0	0	0	0	6
SPAIN	10	5	4	3	1	0	0	0	0	26
FRANCE	11	3	3	3	1	0	0	1	1	40
UK	18	4	1	0	0	0	0	0	0	6
GREECE	8	9	1	2	1	2	0	0	0	31
IRELAND	16	2	5	0	0	0	0	0	0	12
ITALY	5	4	2	4	3	2	1	2	0	62
LUXEMBOURG	11	0	1	0	0	0	0	0	0	2
NETHERLANDS	17	4	1	1	0	0	0	0	0	9
PORTUGAL	20	2	1	0	0	0	0	0	0	4
SWEDEN	15	0	0	0	0	0	0	0	0	0

Note: Trade-related infringement disputes only, 1978-1999

The second explanatory variable is the domestic political constraints in the defendant country, including the institutional checks and balances in the country and the prevalence of partisan divisions. Because this is a cross-national and temporal analysis, I use a composite metric that describes the VETO POINTS in each country and each year (Tsebelis, 1995, 2003; Henisz, 2002). Veto points range from zero to one with a higher number denoting more constraints. Denmark has the fewest while Belgium and the Netherlands have many veto points. Most countries fall into the moderate range (0.4 to 0.6). The veto points variable comes from the Political Constraints Index Dataset. Because of their wide use in closely-related research, these data are appropriate for the present analysis.

I focus on the interaction between adverse trade-related ECJ ruling and domestic veto points which I expect to be negatively correlated with import shares. Even when adverse ECJ rulings have a positive direct effect on subsequent import shares, veto points should diminish this effect, leading to an overall negative correlation. The interaction term is denoted $ATR \times VPs_{t-n-1}$. Note that n denotes the length of the lag in years. This reflects two sources of temporal delays. First, there is the delay between when the adverse ruling is issued (t) and when the defendant government implements the ruling. Although formally the defendant government is obligated to immediately implement rulings, in actual practice, many governments are slow to take action.¹⁶ The additional year of lag reflects the sequence of events where a ruling occurs before the domestic veto players are activated. Second, there is a delay between when a government implements an infringement ruling through law or regulation and when we might observe an economic impact. The multi-year

¹⁶In a particularly extended dispute, France was condemned in a 2004 ruling by the European Court of Justice for failing to implement a directive on the management of waste water. But by 2008, the government still had not taken action. As one news source reported: “although France told the Commission in May last year of plans to install wastewater treatment, the facilities will not be completed until 2011, seven years after the court ruling and 12 years after the directive’s deadline. The Commission said the delay was deplorable. ” “WaterFrance faces court fine over waste delays,” *Utility Week*, February 8, 2008.

lag accounts for these delays. In Section 3.5, a three-year lag is used in all model specifications unless otherwise noted.

3.3.3 Control Variables

I include control variables that may influence a country's imports. These are the log of gross domestic product GDP (measured in US 2005 dollars), GDP GROWTH, GDP PER CAPITA, and the UNEMPLOYMENT rate for the total labor force, controls that are important because countries that are more affluent and enjoy economic growth tend to trade more. Some specifications include sector-specific controls: AGRICULTURE value added, measured as a share of GDP, INDUSTRY value added (as a share of GDP) and MANUFACTURING value added (as a share of GDP). Because some sectors may be more responsive to policy changes and others take longer to adapt, controlling for these differences is important. The control variables come from the World Development Indicators database (World Bank, 2013).

In addition, I aim to control for temporal trends that potentially affect import shares. In some specifications, I include dummy variables for the two structural changes in 1995 and 1999 when the EU expanded membership significantly EU ENLARGEMENT and when countries first adopted the EUROZONE. Other specifications include a time trend variable. All specifications contain a lagged dependent variable to control for other unobserved factors.

Because the VETO POINTS metric is a composite measurement of institutional and partisan factors, it is useful to control for other features of domestic government that may be related to EU integration. Specifically, I control for the defendant government's FRACTIONALIZATION,¹⁷ political POLARIZATION,¹⁸ and an indicator

¹⁷Fractionalization measures the probability that two random draws would produce legislators from different political parties.

¹⁸Political polarization, measured 0 to 2, is the maximum difference of political orientation among the four major parties in government.

for whether the executive’s political party is right, center or left, denoted EXECUTIVE CENTER and EXECUTIVE LEFT. These data come from the Database of Political Institutions (Beck et al., 2001; Keefer and Stasavage, 2003).

3.3.4 Model Estimation

In the sections that follow, I analyze these data using classical (frequentist) and Bayesian statistics. In the classical analysis, I employ the following baseline model specification:

$$\begin{aligned} \text{ImportShare}_t = & \alpha + \beta_1 \text{ImportShare}_{t-n} + \beta_2 \text{Rulings}_{t-n-1} + \beta_3 \text{VetoPoints}_{t-n} \\ & + \beta_4 \text{Rulings}_{t-n-1} \times \text{VetoPoints}_{t-n} + \beta_5 X_{t-n} + \dots + \epsilon \end{aligned}$$

The regression equation includes a lagged dependent variable (n years) to ensure estimates reflect deviations from a baseline trade share. By controlling for initial import share, I am able to examine how each member’s economic integration has changed in the wake of ECJ rulings. The β coefficients estimate the effect of domestic veto points, the number of adverse ECJ trade rulings in a year, and their interaction. The control variables are denoted by X . In various model specifications, I adjust for temporal autocorrelation with an AR(1) correction and panel corrected standard errors.

One challenge is identifying the extent to which country-level variation versus temporal variation drives the estimates. To investigate this, I introduce hierarchical structure to the data, modeling country-random intercepts, random slopes for the interaction term, and country-level variables in certain specifications. While multilevel modeling in a frequentist framework is certainly a reasonable option, I use Bayesian analysis here. Bayesian analysis makes hierarchical modeling particularly convenient

and flexible, allowing me greater control over the functional forms and distributions of parameters. I employ the following model specification and variations thereon:

$$\begin{aligned}
y_{it} &\sim N(\alpha_i + \beta_i x, \sigma_y^2) \\
\alpha_i &\sim N(\mu_a, \sigma_a^2) \\
\beta_i &\sim N(\mu_b, \Omega) \\
\sigma_y, \sigma_a &\sim \text{Unif}(0, 100) \\
\mu_a, \mu_b &\sim N(0, 100) \\
\Omega &= \begin{bmatrix} \omega^2 & 0 \\ 0 & \omega^2 \end{bmatrix}
\end{aligned}$$

The beta coefficients are given flat priors with variance $\omega^2 = 100$ and assumed to be uncorrelated with one another. The covariance matrix Ω is shown for two betas for simplicity and is extended to a K dimensional diagonal matrix where K is the number of covariates in the model.

3.4 Results - Classical Approach

The classical (frequentist) analysis lends support to the theory. Countries that lose ECJ infringement lawsuits tend to increase their EU import share in subsequent years but this tendency is mitigated by domestic politics. Adverse rulings are correlated with increasing intra-EU imports. But countries that lose ECJ lawsuits and have many veto points import less from the EU in subsequent years. As shown in Table 3.2, the *interaction* between domestic veto points and adverse ECJ trade rulings is negatively related to trade. I interpret this to mean that when countries have more domestic political constraints, they are less responsive to adverse ECJ rulings on trade issues—their import share is lower than would otherwise be expected. The

correlations become more substantial as the time lag is increased from one to three years, allowing more time to elapse between the ruling and economic outcomes. The temporal lags are appropriate since most legal disputes at the European Court of Justice take at least two years to resolve and there are often further delays before the effect of a policy is felt on commerce.

Because the regressions in Table 3.2 show evidence of heteroskedasticity and temporal autocorrelation, I include additional corrections. Table 3.3 presents results for models with AR(1) Prais-Winsten corrections and panel-corrected standard errors. The TIME variable accounts for temporal trends in the data. Some specifications include sectoral controls. Countries whose economies rely heavily on manufacturing import relatively less whereas countries that rely on agriculture and industry import relatively more. The intercept is omitted in these model specifications to avoid “overfitting.” These estimates suggest that in the face of a legal intervention from the ECJ, domestic constraints appear to hinder a country’s ability to respond and implement policies that expand intra-EU commerce.

I also consider certain aspects of domestic government that may be associated with EU integration. Table 3.4 presents several such variables. Defendant governments with highly polarized legislatures are less likely to see their intra-EU import shares increase in the wake of ECJ rulings. Governments led by centrist parties—Ireland, Italy, and Luxembourg in various years—tend to have higher EU import shares than right-leaning governments. Left-leaning governments—which nearly all EU countries feature at some point in time—tend to have somewhat lower import shares in the following years.

Table 3.2: Linear Models for EU-14 Import Share with Various Lags, 1988-2005

	EU Import Share $n = 1$ year (<i>OLS</i>)		EU Import Share $n = 2$ years (<i>OLS</i>)		EU1 Import Share $n = 3$ years (<i>OLS</i>)	
	(1)	(2)	(3)	(4)	(5)	(6)
Share $_{t-n}$	0.928*** (0.060)	0.891*** (0.072)	0.923*** (0.083)	0.971*** (0.108)	0.772*** (0.085)	0.796*** (0.101)
ATR \times VPs $_{t-n-1}$	-0.021 (0.026)	-0.016 (0.029)	-0.079* (0.042)	-0.090** (0.042)	-0.111*** (0.040)	-0.118*** (0.037)
ATR $_{t-n-1}$	0.016 (0.017)	0.015 (0.018)	0.045* (0.027)	0.051* (0.027)	0.075*** (0.025)	0.079*** (0.024)
Veto Points $_{t-n}$	-0.001 (0.019)	0.0001 (0.018)	0.0002 (0.028)	0.002 (0.030)	-0.007 (0.028)	-0.005 (0.029)
GDP Growth $_{t-n}$	-0.001 (0.002)	-0.002 (0.002)	0.006* (0.003)	0.007** (0.003)	0.011*** (0.004)	0.012*** (0.004)
GDP $_{t-n}$		-0.003* (0.001)		0.005 (0.003)		0.003 (0.004)
Unemployment $_{t-n}$		-0.0004 (0.001)		-0.001 (0.001)		-0.001 (0.001)
EU Enlargement	-0.025* (0.014)	-0.032* (0.016)	-0.047** (0.022)	-0.037 (0.026)	-0.082*** (0.023)	-0.077*** (0.026)
EuroZone	0.031*** (0.010)	0.024** (0.012)	0.076*** (0.016)	0.081*** (0.019)	0.090*** (0.016)	0.091*** (0.019)
Constant	0.013 (0.038)	0.039 (0.045)	-0.032 (0.052)	-0.054 (0.067)	0.008 (0.052)	0.001 (0.066)
Observations	246	246	246	246	246	246
# Countries	14	14	14	14	14	14
R ²	0.898	0.898	0.843	0.844	0.814	0.815

Note: Regressions of interaction between adverse ECJ rulings for trade-related infringement disputes and domestic veto points. Panel-corrected standard errors. Significance codes:

*p<0.1; **p<0.05; ***p<0.01

Table 3.3: Linear Models for EU-14 Import Share with Various Lags and Economic Controls and AR(1) Correction, 1988-2005

	EU Import Share $n = 1$ year (<i>OLS</i>)		EU Import Share $n = 2$ years (<i>OLS</i>)		EU Import Share $n = 3$ years (<i>OLS</i>)	
	(1)	(2)	(3)	(4)	(5)	(6)
Share $_{t-n}$	0.932*** (0.041)	0.859*** (0.065)	0.788*** (0.077)	0.547*** (0.119)	0.650*** (0.089)	0.378*** (0.126)
ATR \times VPs $_{t-n-1}$	-0.060** (0.030)	-0.054* (0.031)	-0.155** (0.060)	-0.145*** (0.053)	-0.221*** (0.070)	-0.167*** (0.060)
ATR $_{t-n-1}$	0.027 (0.017)	0.030* (0.018)	0.072** (0.028)	0.071*** (0.025)	0.130*** (0.031)	0.112*** (0.026)
Veto Points $_{t-n}$	0.042* (0.024)	0.029 (0.018)	0.121** (0.050)	0.092*** (0.034)	0.133** (0.060)	0.082** (0.037)
GDP $_{t-n}$	-0.001 (0.002)	-0.001 (0.003)	-0.001 (0.000)	-0.001 (0.006)	-0.004 (0.006)	-0.005 (0.008)
Unemployment $_{t-n}$	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.002)	-0.002 (0.002)	0.000 (0.002)	-0.002 (0.002)
GDP Growth $_{t-n}$	-0.001 (0.003)		0.004 (0.005)		0.009* (0.005)	
Agriculture VA $_{t-n}$		0.003 (0.003)		0.012** (0.005)		0.017*** (0.006)
Industry VA $_{t-n}$		0.003** (0.001)		0.011*** (0.003)		0.013*** (0.004)
Manufacture VA $_{t-n}$		-0.003** (0.001)		-0.009*** (0.003)		-0.011*** (0.004)
Time	-0.000 (0.001)	-0.001 (0.001)	-0.002 (0.002)	-0.005* (0.003)	-0.002 (0.002)	-0.005 (0.003)
Observations	276	241	261	227	246	213
# Countries	14	14	14	14	14	14
R ²	0.914	0.904	0.753	0.756	0.682	0.715

Note: Regressions of interaction between adverse ECJ rulings for trade-related infringement disputes and domestic veto points. Panel regression with AR(1) Prais-Winsten correction and panel-corrected standard errors. Intercept omitted due to time index. Significance codes:

*p<0.1; **p<0.05; ***p<0.01.

Table 3.4: Linear Models for EU-14 Import Share with Political Controls, 1988-2005

	<i>Dependent variable:</i>				
	EU Import Share				
	$n = 3$ years (<i>OLS</i>)				
	(1)	(2)	(3)	(4)	(5)
Share _{$t-n$}	0.408*** (0.067)	0.397*** (0.068)	0.377*** (0.066)	0.389*** (0.067)	0.322*** (0.069)
ATR×VPs _{$t-n-1$}	-0.050** (0.022)	-0.038* (0.022)	-0.046** (0.022)	-0.050** (0.022)	-0.040* (0.022)
ATR _{$t-n-1$}	0.011 (0.011)	0.006 (0.011)	0.010 (0.011)	0.011 (0.011)	0.008 (0.011)
Veto Points _{$t-n$}	0.168* (0.090)	0.196** (0.098)	0.149* (0.089)	0.149 (0.092)	0.085 (0.100)
GDP _{$t-n$}	0.120* (0.070)	0.118* (0.071)	0.120* (0.068)	0.103 (0.071)	0.037 (0.072)
Fractionalization _{$t-n$}		-0.111 (0.120)			0.048 (0.149)
Polarization _{$t-n$}			-0.026*** (0.009)		-0.026** (0.010)
Center Executive _{$t-n$}				0.023 (0.028)	0.070** (0.031)
Left Executive _{$t-n$}				-0.016 (0.011)	-0.019* (0.011)
Time	-0.015*** (0.002)	-0.015*** (0.002)	-0.015*** (0.002)	-0.014*** (0.002)	-0.014*** (0.002)
Observations	252	242	252	252	242
# Countries	14	14	14	14	14
R ²	0.775	0.773	0.784	0.779	0.789

Note: Regressions of interaction between adverse ECJ rulings for trade-related infringement disputes and domestic veto points. Panel regression panel-corrected standard errors. Intercept omitted due to time index. Significance codes: *p<0.1; **p<0.05; ***p<0.01.

3.5 Results - Bayesian Approach

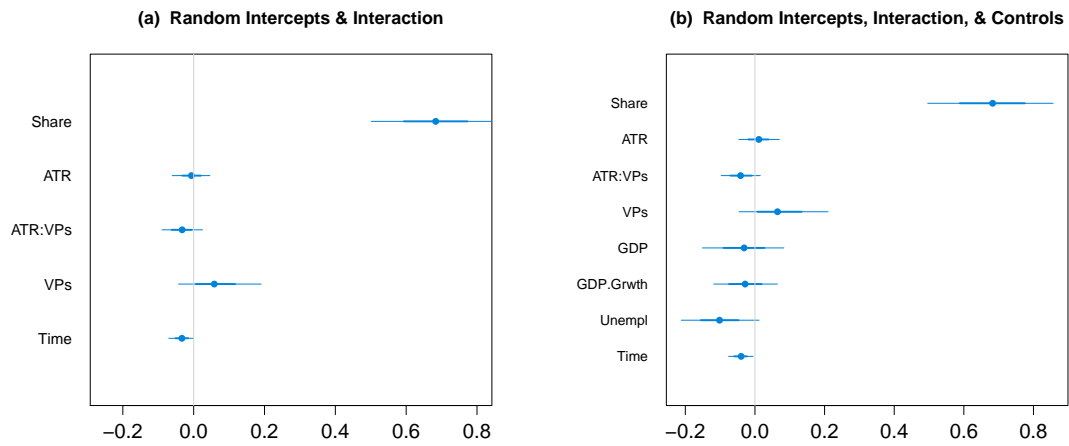
3.5.1 Random Intercepts Model

One concern in the preceding analysis is whether there are heterogeneous effects across countries. Are certain countries driving the observed trends? A Bayesian approach allows for flexible modeling of the hierarchical structure in the data (Jackman, 2009) which gives me greater leverage on these issues. I begin by examining the estimated effect of adverse ECJ rulings and domestic veto points on intra-EU trade. This analysis confirms the main substantive findings above. First, adverse ECJ rulings on trade are positively correlated with a country's EU import share, controlling for the initial import share (lagged dependent variable). Second, the interaction effect of an adverse ECJ rulings and domestic veto points is negative. This tendency is statistically significant when economic control variables and a time index are included in the model.

To evaluate heterogeneity across countries, I use a random intercepts model where a separate intercept is fit to each country. The intercepts are assumed to be distributed normally and I use uninformative “flat” priors. The model is implemented with **Bugs** code provided in the Appendix B.¹⁹ I test two random intercepts model specifications whose coefficients are displayed in Figure 3.2. The simpler model (left) includes the interaction between domestic veto points and adverse ECJ rulings relating to trade. The more complex model (right) further includes economic controls. In both, the interaction term is negative and nearly significantly different from zero. The latter specification will be the focus of the remaining analysis.

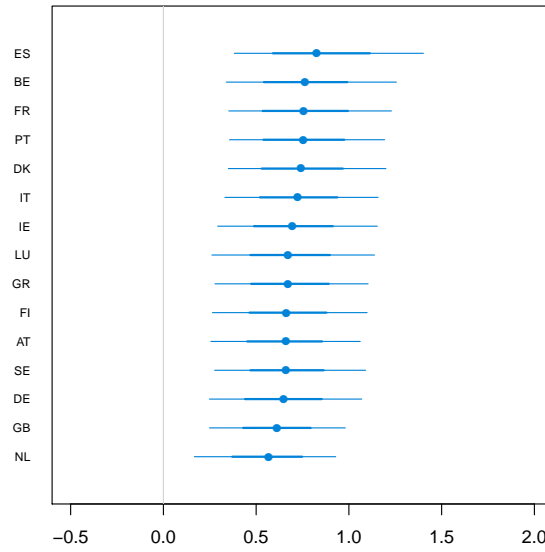
¹⁹All continuous variables are normalized before entering the analysis. That the prior precision for the distribution of beta coefficients is specified outside the **Bugs** code as a diagonal matrix of 0.001. Each model is run with two chains of length 50,000 with thinning.

Figure 3.2: Parameter Estimates for Country-Random Intercept Models



Following the standard assumption of normally distributed random intercepts, if one country's posterior intercept lies outside the distribution or the intercepts are not normally distributed, this would indicate the specification is not modeling the main structure of the data. Figure 3.3 shows the country-random intercepts are roughly normally distributed (with a wider tailed distribution than standard normal). No outliers are evident, suggesting that the model fits each country equally well.

Figure 3.3: Country-Random Intercepts

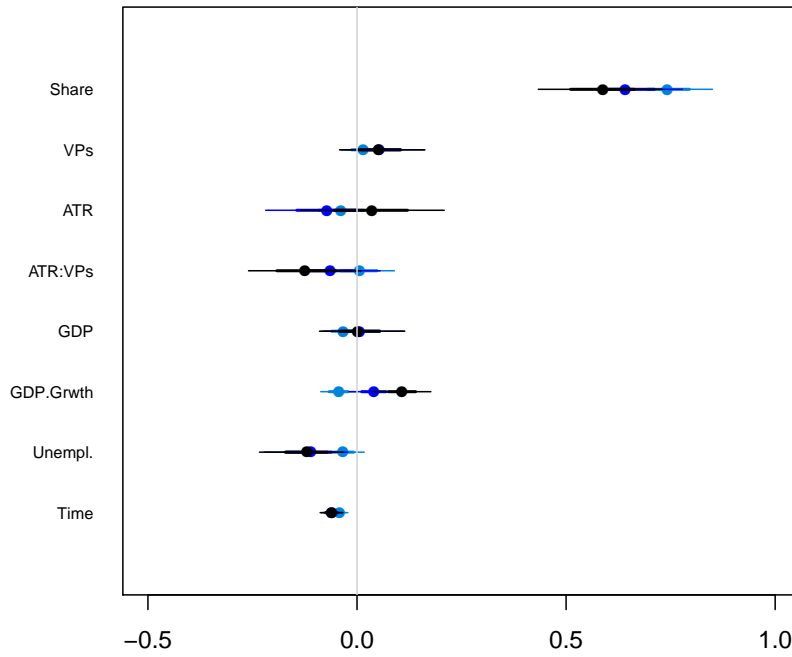


Note: Intercepts correspond to interaction model with controls, as shown in Figure 3.2 (b).

The estimates are fairly robust to variation in the temporal lag structure. The models above include a three-year lag on all explanatory variables and controls except the adverse ECJ ruling (ATR_{t-n-1}) which is lagged an additional year. As a robustness check, I repeat the analysis using one- and two- year lags. Figure 3.4 compares the results (light blue denotes a one-year lag, dark blue denotes a two-year lag, and black denotes the three-year lag presented above.) The size of the estimated effect of the interaction term, $ATR \times VPs_{t-n-1}$ varies. A longer lag corresponds to a larger estimated effect. This result may reflect an actual effect of domestic veto points and adverse rulings on trade—time allows the impact of policy on trade to be borne out. Or, this result may merely be a statistical artifact of using a lagged dependent variable.²⁰

²⁰The longer the lag, the less precisely the lagged dependent variable is correlated with the dependent variable, leaving more variance in the data to be explained by the other covariates.

Figure 3.4: Parameter Estimates Varying Temporal Lag



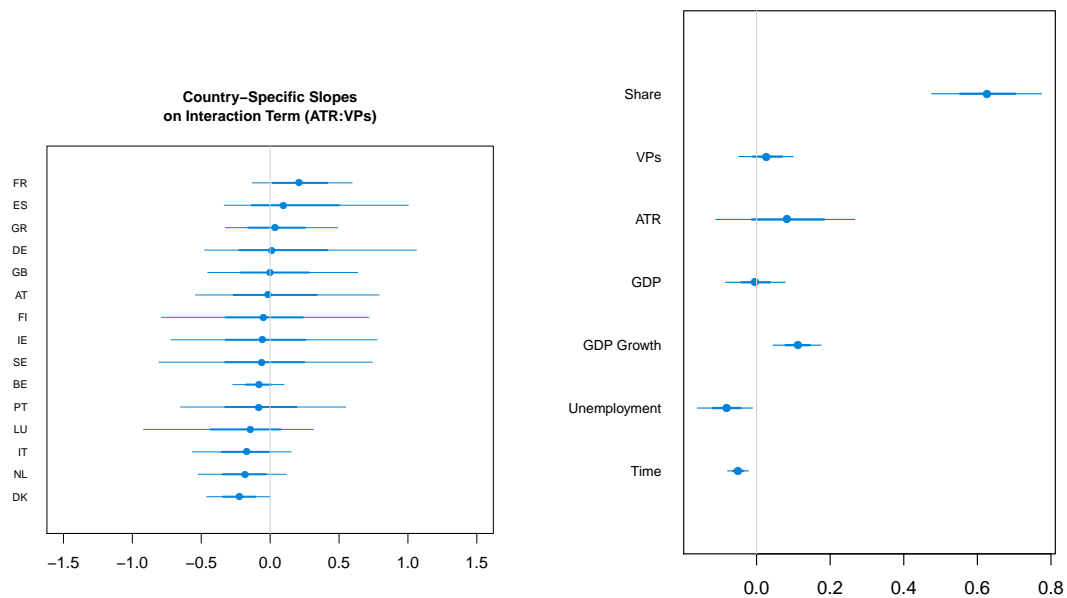
Note: Estimates are for the interaction model with controls and country random intercepts. Light blue denotes a one-year lag, dark blue denotes a two-year lag, and black denotes the three-year lag.

3.5.2 Random Slopes Model

The interaction term between domestic veto points and adverse ECJ rulings, the focal point of the analysis, might also be sensitive to influential countries. To assess this possibility, I repeat the analysis using a country-random slopes model. Each country receives a different coefficient on the interaction term and those coefficients are assumed to be normally distributed. Figure 3.5 displays coefficient estimates for country-random slopes model specification. The estimates for the interaction term $ATR \times VPs_{t-n-1}$ are negative but not significantly different from zero for most countries (left panel). Denmark and the Netherlands display the most negative

slopes while France and Spain have the most positive slopes. The coefficients for other covariates are generally consistent with previous estimates (right panel). The unemployment rate is correlated with decreased trade shares while GDP growth is correlated with increased trade shares. These results are substantively plausible since an economic downturn and improvement should be linked to expanding and shrinking intra-EU trade, respectively. These models use a single intercept to reduce the number of parameters.

Figure 3.5: Country-Random Slopes on Interaction Term and Shared Parameter Estimates for Other Covariates



Note: Left panel shows country-random slopes on interaction term and right panel shows shared coefficient estimates for other covariates.

I also fitted a model with both country-intercepts and country-slopes for the interaction term and found the estimates to be quite similar. Estimates for the interaction term are generally negative and not significant.

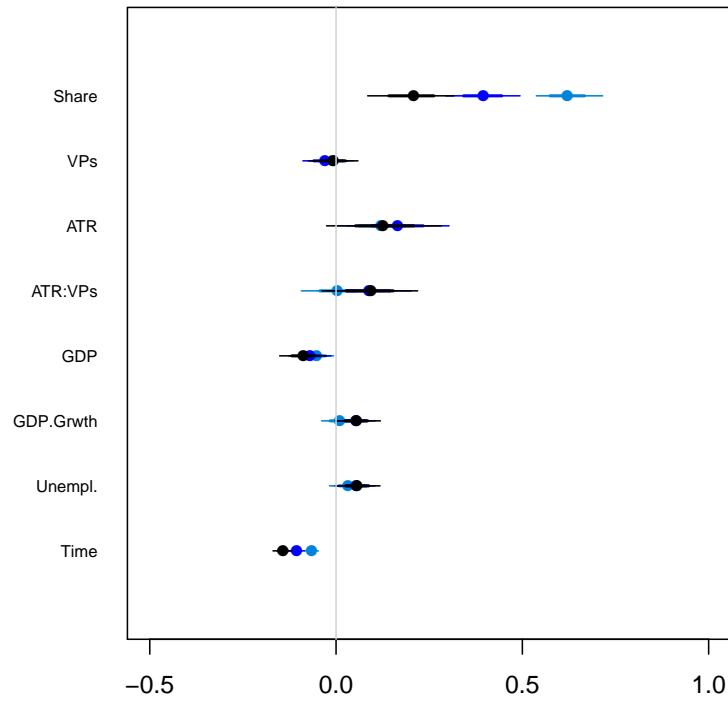
3.5.3 Robustness

3.5.3.1 Placebo Test

One difficulty in interpreting these results is that spurious correlations in the data may be driving the findings. If spurious correlations are the culprit, the direction of causality I attempt to model should not substantially change the estimated coefficients. Each country's EU import share is fairly steady over time so for a given country, trade in one year is highly correlated with trade in any other year. So then spurious correlations should produce estimates where an adverse ECJ ruling and domestic veto points correspond to lower trade shares in the preceding one, two, and three years. In other words, finding the same estimated "effect" in the opposite direction would undermine the proposed mechanism.

To assess this, I conduct placebo tests for reverse causality. Instead of lagging the explanatory variables, I advance them by one, two, and three years and re-run the baseline model (without random intercepts or slopes). The estimated coefficients are presented in Figure 3.6 where light blue denotes a one-year lead, dark blue a two-year lead, and black a three-year lead. In all cases, the estimated effect of the interaction term $ATR \times VPS_{t-n-1}$ is positive and not significantly different from zero. This null effect lends support to the findings presented above.

Figure 3.6: Coefficient Estimates for Placebo Test



Note: Estimates are for the interaction model with controls and country random intercepts. Light blue denotes a one-year lead, dark blue denotes a two-year lead, and black denotes the three-year lead.

3.5.3.2 Influential Country Test

Finally, I assess whether one or more influential countries are driving the findings by repeating the baseline Bayesian analysis for subsets of the data. In each subset, observations for a single country are removed. I calculate the estimated effects and by comparing to estimates with the full sample, assess the sensitivity of the coefficient estimates to influential countries. Generally, the coefficients are approximately normally distributed, indicating that no single country tends to drive the estimates. However as noted above in Figure 3.5, the interaction term estimate is sensitive to the inclusion of Denmark which is especially influential in generating the negative

correlation. Denmark also appears influential in estimating the coefficient on the unemployment rate variable.

3.6 Discussion

The European Court of Justice is widely recognized as an effective and successful international court. Some scholars have argued that the ECJ was used as a template for other international legal regimes (Alter and Helfer, 2010; Alter, Helfer, and Saldias, 2012). Part of this success can be attributed to the ability of the court to supersede national political strife, avoiding the detrimental effects of narrow interest group lobbying that might otherwise undermine EU cooperation.

With its doctrines of supremacy and direct effect, the ECJ has strong legal basis for automatic implementation of international rulings in domestic law and indeed has been successful in doing so in many cases. Supremacy means that where there is an irresolvable conflict between national law and European law, it is the European legal obligation that must be applied—it supersedes national law. Direct effect means that ECJ rulings can be invoked and enforced in domestic court and therefore are more likely to lead to compliance with rulings. Relying on these principals, the preliminary reference system appears to have been quite effective in linking international rulings to domestically binding law (Carrubba and Murrah, 2005). The preliminary reference procedure is embedded in the domestic legal system of the member states. It allows national judges to refer questions of EU law to the ECJ for a ruling on matters relating to EU law and requires they apply the ECJ’s decision in the final disposition of the national case. Gabel et al. (2012) have found a positive effect of preliminary rulings on intra-EU trade—but not of infringement disputes.

For infringement cases, which are decided purely at the international level and implemented separately, at the domestic level, ECJ rulings are less embedded into the national legal systems. There is indeed a tendency for domestic politics to

interfere. Where the national government shoulders the burden to take deliberate actions and change domestic policy to comply with ECJ rulings, politics can be a hindrance. While previous scholars have argued that bureaucratic capacity is a barrier, with governments that want to comply but simply lack the administrative ability to put ECJ rulings into effect, these findings point to a different obstacle. Rather, governments encounter opposition from groups that stand to lose from EU integration. When governments have many veto players, those opposition groups have inroads to block policy reforms and thereby derail compliance with ECJ rulings. Strong domestic opposition groups can undermine EU integration.

This analysis has demonstrated a consistent link between adverse rulings from the European Court of Justice, domestic politics, and European economic integration. Countries that face adverse judicial rulings on trade-related issues subsequently expand their intra-European imports, relative to their total imports from the world but domestic political constraints appear to mitigate this response. The *interaction* between adverse ECJ ruling and domestic political divisions is negatively and significantly correlated with intra-EU import shares in the years following the ruling. The empirical findings suggest adverse rulings from the European Court of Justice can prompt defendant countries to open their markets to more European commerce, but that domestic political constraints hinder the process. Domestic constraints, measured as veto points, may inhibit the policy changes mandated by the ECJ and thereby obstruct European economic integration.

Two methodological approaches demonstrate this finding. First, using classical (frequentist) statistics, I show that the correlation is strong when controlling for other important economic factors and when the data are modeled with different temporal lags and error structures. Second, using a Bayesian approach, I introduce more hierarchical structure. This analysis shows that the estimated effects are somewhat heterogeneous across countries but are ultimately quite robust. I show that the

estimated effects are unlikely to arise purely from spurious correlations in the data; reversing the temporal sequence in the model does not generate the same estimates.

The findings speak to a broader literature on the effectiveness of international courts in promoting economic cooperation among countries and how domestic politics matter. With strong institutions and more extensive enforcement capabilities than most international institutions, the EU might be perceived as especially insensitive to domestic discord. The small but statistically reliable impact of veto points on post-ruling EU imports suggests that even here, domestic political discord can have deleterious effects on compliance. Supremacy, direct effect, and enforcement mechanisms may not be enough to transform the ECJ into the binding judicial authority it is often represented to be. While the European Court of Justice is often recognized as a highly effective international court, its powers remain limited (also see: Carrubba (2005)). For infringement disputes, European countries remain reluctant compliers. Institutional checks and partisan divisions at the domestic level can obstruct the implementation process, stalling economic integration.

CHAPTER 4

Comparing International Courts: How Institutional Design Mediates the Domestic Politics of Dispute Settlement

4.1 Introduction

As long as international courts have issued rulings with domestic political significance, they have provoked push-back from national governments. Politicians are more than willing to publicly declare their disapproval of international courts and their verdicts. For example, in 2008 in a strongly-worded editorial, former German president Roman Herzog condemned the European Court of Justice for “deliberately and systematically ignoring fundamental principles of the Western interpretation of law...undermining the core competencies of the member states.”¹ In 1999, amidst widespread protestation of a World Trade Organization legal ruling striking down a US environmental policy, Bill Clinton referred to the WTO as a “private priesthood for experts” and vowed to try and open the organization to environmental groups.² Even the newest major entrant to the WTO, China, is reluctant to defer to the institution’s legal rulings. Concerning a dispute over tariffs on automobile parts, the Chinese trade official Zhou Shijian stated that the adverse ruling was a “groundless accusation,” adding “the regulations of the WTO are made by the developed coun-

¹“Stop the European Court of Justice,” *EU Observer*, September 10, 2008.

²“Clinton Seeks Seat for Critics at WTO,” *The Seattle Times*, October 14, 1999.

tries, after all.”³ And yet governments continue to invest heavily in legalized dispute settlement, buttressing the institutional authority of these courts and introducing enforcement mechanisms that increase the likelihood that rulings are implemented. Sometimes governments comply with adverse rulings from international courts; sometimes domestic political obstacles are insurmountable. What explains the different outcomes? In part, the answer depends on the court’s design, with more flexible institutional designs giving governments greater leeway to accommodate the demands of domestic politics.

This chapter presents a theory of the domestic politics of compliance with international courts and how the institutional design of the court matters. I argue that the domestic political constraints in government—veto players—can impede compliance by reducing a government’s ability to conform to the legal rulings from international courts. When the court is flexible, in that it allows countries opportunities to prolong the legal process and imposes minor penalties on recalcitrant countries, the effect of domestic veto players is most pronounced. Because these courts allow countries substantial leeway, defendant governments can accommodate the demands of domestic political groups that oppose compliance while avoiding international punishment. Conversely when the court is rigid, in that it permits countries few means to deviate from their obligations and is able to penalize noncompliance, the effect of domestic veto players is modest. The design of the institution mediates the link between domestic politics and international cooperation.

To evaluate this theory, I compare two prominent international courts. The first is the World Trade Organization (WTO) Dispute Settlement Mechanism. It is a relatively flexible and permissive legal institution that allows countries room to negotiate, compensate, and, in extenuating circumstances, permits enforcement by plaintiff governments. It does not have a central institutional mechanism to ensure

³“First Six Years of China in WTO Uncertain,” *Oriental Morning Post*, February 15, 2008, available <http://auto.sina.com.cn/news/2008-02-15/0803347384.shtml>.

countries comply with its rulings. The second is the European Court of Justice (ECJ). As compared to the WTO, the European Court of Justice is a relatively rigid institution with strict legal procedure, enforcement capacity, and little opportunity for countries to negotiate. Because both courts adjudicate fundamental aspects of international economic cooperation and have the most active track records in international law, they are ripe for comparison. Yet as the literatures on European law and the WTO have become increasingly specialized, few studies have endeavored to bridge the two. Using institutional comparisons, this study aims to lend greater insight into the domestic politics of international dispute settlement.

First, I show that the two institutions produce divergent outcomes: disputes at the WTO last longer than ECJ disputes. The dispute period includes both litigation and, in the case of an adverse ruling, the period in which a losing defendant must adjust its policy. Losing defendants at the WTO have incentive to prolong noncompliance and resist the court's authority as long as possible. By contrast, losing defendants at the ECJ have incentives to comply promptly and avoid penalty. In both courts, defendants sometimes defy the first legal ruling and face another round of litigation, further increasing dispute duration. This is more common at the WTO than the ECJ, where losing defendants are rarely accused of continued noncompliance.

Second, I show that disputes are more difficult to resolve when domestic politics is divisive. When a defendant government has many veto players it prolongs litigation and, when faced with an adverse ruling, is more likely to undergo noncompliance proceedings. The link between domestic veto players and international behavior is pronounced under the WTO but negligible under the ECJ. These empirical patterns, along with detailed examples, supply evidence supporting my theory of international courts and the domestic politics of compliance.

The results speak to a broader debate about the effectiveness of international institutions. Countries generally appear to comply with their international com-

mitments, and many scholars have interpreted this as evidence that international institutions work (Chayes and Chayes, 1993; Simmons, 1998). But because countries voluntarily join international institutions, they are likely to select treaty terms that they intend to follow anyway, even in the absence of the institution. Widespread compliance may be better explained by the selection process rather than the binding force of international agreements. As a result, it is problematic to draw inferences about institutional effectiveness from apparent compliance rates (Downs, Rocke, and Barsoom, 1996; Von Stein, 2005). By examining international legal disputes, we gain some leverage on this inferential puzzle. When an international court rules against a defendant government, it determines that the country has violated its international commitments. In these instances, it is reasonable to assume the defendant government preferred the policy or action that violates the international agreement. If the country then corrects the violation, complying with the international ruling despite its preference, we can infer that the court promoted compliance. Countries' response to adverse rulings provides a useful vantage for evaluating the effectiveness of international courts.

In the section that follows, I discuss the WTO Dispute Settlement Mechanism (DSM) and European Court of Justice, highlighting institutional design features that are thought to impact compliance. Next, I present a theoretical framework linking domestic politics to dispute settlement, explaining how the design of the international court can mediate this relationship. In the fourth section, I provide empirical tests using data on legal disputes at both international courts. In the last section, I discuss the findings and implications for international cooperation more broadly.

4.2 Comparing International Courts

4.2.1 International Courts

The international economic landscape is becoming increasingly legalized (Goldstein and Martin, 2000). As trade flows and investment across national borders become major parts of countries' economies, governments have sought ways to regulate them through international law. The multilateral trade regime, embodied by the WTO, manages most major aspects of global trade relations.⁴ The European Union now regulates economic relations among European countries and between Europe and the rest of the world. These international economic institutions aim to facilitate cooperation not only through legal standards, but also with bureaucratic bodies that support countries' implementation of treaty terms through technical assistance, monitoring, and other means.

The two institutions are similar in numerous ways that make them suitable for comparison. Both restrict the trade policies of member countries and create deep legal obligations. The depth of the agreement is defined by the extent to which it requires reduction of barriers to trade, behind-the-border integration with regard to social and environmental standards, or otherwise imposes stringent constraints on the policies available to the member states. It is thought to be a fundamental element of institutional design (Gilligan, 2004).

The European Union (EU) imposes deep obligations on its members in several ways. Its agreements create a common internal European market where goods and services can move freely across national borders, a common external tariff and tight restrictions on the use of non-tariff barriers, strict limits on state subsidies and anticompetitive practices by firms ("competition policy"), a shared currency—the Eurozone—and many other dimensions of economic integration. Due to the depth

⁴Notably, the WTO does not regulate trade related aspects of labor or environmental policy, an issue of much contention.

of EU economic cooperation and the central bureaucracy that serves to execute the agreement terms, some scholars describe the EU as one of the most “ambitious and demanding” of international institutions and others characterize it as a federation (Phelan, 2015; Bednar, 2009).⁵

The WTO also places substantial restrictions on the trade policies of its members. Since 1995 when it replaced its predecessor, the General Agreement on Tariffs and Trade, the WTO has grown to have global reach with 161 members (European Union countries participate as a single member block). Its agreements cover all sorts of tariff and non-tariff barriers to trade pertaining to agriculture, intellectual property, services, and many other areas. Like the European Union, the WTO’s obligations affect fundamental aspects of economic integration, all but eliminating the use of tariffs and constraining governments’ use of subsidies, discriminatory taxes, procurement rules, and other policies with trade implications.

Because both international institutions rely on the voluntary cooperation of member countries, conflicts inevitably arise. Sometimes countries violate their primary legal obligations, either by failing to implement the appropriate policies in the first place or by imposing prohibited trade barriers or other policies. In these instances, the countries have not followed through on their international commitments. When governments are accused, either by other governments or by the international organization itself of violating treaty terms, they often seek resolution through legal means. International courts provide the necessary venue.

Under both institutions, international courts are seen as an important tool for preserving or restoring cooperation between countries. By providing a legal venue for dispute settlement, these courts help member governments reach peaceful resolution according to agreed-upon rules and avert more costly forms of conflict like trade wars where the cooperative equilibrium breaks down. Countries collectively

⁵Note that the European Union began as the European Economic Community in 1958 and was renamed in 1992. For convenience, I use the label “European Union” throughout the entire period.

prefer to participate in legalized dispute settlement over conflict and *ad hoc* bilateral agreements. To this end, international courts are common and central pillars of many international organizations.

The WTO and European Union’s judicial bodies are the most active and (potentially) influential we have yet seen in the economic domain. Both aim to evaluate complaints about alleged treaty violations and determine whether the defendant government has followed through on its legal obligations. If the defendant is found to be in breach of treaty terms, the goal is to bring it into compliance. These judicial bodies are discussed in detail below.

Legal verdicts from an international court will often require the defendant country to revise its policy and bring its practices into conformity with treaty terms. A government that faces an unfavorable legal outcome has a unilateral incentive to ignore the ruling and evade the court. For example, if a government loses a WTO dispute over its use of a safeguard—that is, the DSM determines the trade barrier is not legal—it may want to defy the ruling and keep the safeguard in place. Or in another example, if the ECJ rules against a member for its use of a discriminatory tax, the government may want to resist and retain the policy. By keeping the policy in place, a government may enjoy certain benefits like a boost in political support from the regions or interest groups representing the protected industry. With these domestic pressures, both the DSM and the ECJ confront compliance problems.

4.2.2 The WTO Dispute Settlement Mechanism

The WTO Dispute Settlement Mechanism is central to the successful operation of the multilateral trade regime. The WTO explains, the “dispute settlement system is the backbone of today’s multilateral trading regime...[and] ensures that the WTO’s carefully negotiated trading rules are respected and enforced.”⁶ Since 1995—when

⁶ *WTO Dispute Settlement System Training Module*.

the DSM was established in its current form—to the present, it has presided over 471 trade disputes. With nearly universal membership, jurisdiction over a broad range of trade related policies, and a compliance record that appears strong, DSM is seen as a benchmark for the resolution of international economic disputes.

WTO disputes arise when a plaintiff government sues a defendant government over an alleged breach of treaty terms. Typically, this occurs when the defendant government imposes a trade policy that allegedly harms the plaintiff by either restricting imports into the defendant country or promoting its exports. These trade policies can take many forms including import barriers like safeguards, anti-dumping duties, and countervailing duties as well as export promotion measures like subsidies. The disputes span many industries and topics from health and safety of agricultural goods to the taxation of manufactured goods like semiconductor chips and automotive parts.

In the initial stages, the governments enter formal negotiations to address the alleged violation. Many WTO disputes are settled at this stage with the defendant clarifying the legality of its policy, correcting a violation, or otherwise dispelling of the conflict (42% in my sample). Others linger on the WTO's docket, unresolved. When negotiations fail, the plaintiff initiates litigation by requesting a panel of judges to review the merits of its complaint and issue a legal ruling, a “panel report.” The vast majority of these reports favor the plaintiff on at least one legal claim and thereby confirm that the defendant government has violated its treaty obligations. At this point, many disputants appeal the ruling and go through additional legal proceedings before an “appellate body” which may modify the panel report.

The WTO expects defendants to implement adverse rulings in a timely manner. The defendant is obligated to correct the violation promptly by revoking or revising the trade policy that was in breach of its WTO rules. In actuality, many defendants cannot or are unwilling to comply immediately and the defendants agree upon

a reasonable period of time (RPT) for achieving compliance.⁷ This grace period, typically six to eighteen months, gives the defendant time to amend its domestic law in order to achieve implementation. During that period, the defendant continues to apply its WTO-inconsistent measures without penalty. In the case of continued violations after the grace period, a defendant that has not implemented the ruling faces consequences (i.e. the need to offer compensation or face retaliation). Before the penalties can be imposed, the disputants go through another legal stage called “compliance proceedings” which establish guilt and specify the recourse. The WTO states that compensation or retaliation, through the suspension of concessions, is not preferred to full implementation of the ruling.⁸

Prompt dispute resolution is an explicit goal but it is not consistently achieved. The WTO explains, “justice must not only provide an equitable outcome but also be swift...If a case is adjudicated, it should normally take no more than one year for a panel ruling and...16 months if the case is appealed.” Despite this, some disputes drag on for many years as both governments “dig in their heels.” One of the longest-lasting disputes, *Canada v. European Communities—Hormones*, stretched over nearly 15 years.⁹ Other conflicts give rise to a series of sequential formal disputes, each rehashing the same underlying conflicts between countries. One example is the notorious series of bananas disputes which began in 1995 and were finally resolved in 2012.¹⁰ Dispute settlement at the WTO does not consistently lead to compliance and is often hampered by prolonged delays. The many institutional avenues for delay

⁷WTO Dispute Settlement Understanding, Article 21.3

⁸WTO Dispute Settlement Understanding, Articles 3.7 and 22.1.

⁹This dispute is discussed at length in Section 4.4.

¹⁰DS16, DS27, DS105, and DS158: European Union – Regime for the Importation, Sale and Distribution of Bananas; DS165: United States – Import Measure on Certain Products from the European Communities; DS361 and DS364: European Union – Regime for the Importation of Bananas.

introduce flexibility into the WTO's system because countries can use legal means to prolong noncompliance and cater to the domestic pressures they face.

4.2.3 The European Court of Justice

Like the WTO's Dispute Settlement Mechanism, the European Court of Justice is central to the operation of the European Union. Its purpose is to interpret the burgeoning body of EU law and ensure it is applied uniformly across all member countries. It handles different types of cases including: (1) infringement disputes, where the Court evaluates whether a member state has fulfilled its obligations under EU law and (2) references for preliminary rulings, where national courts request a binding interpretation of EU law. This study focuses on infringement disputes.

Infringement disputes at the ECJ are a suitable venue to examine the impact of domestic politics and draw comparisons to the WTO. Infringement disputes occur when the European Commission, the institution's central administrative body, sues a member state for failing to implement EU law. Typically, the state is aware of its EU obligations, urged to follow through on them, and yet resists. The ECJ rulings on these disputes must then be implemented through the domestic policy process. There is clear separation between the international ruling by the court and compliance with that ruling, which must be taken by domestic actors. ECJ preliminary references, on the other hand, are not particularly comparable to WTO disputes. Preliminary references arise when member states have lawsuits in their domestic courts that hinge on EU law and are (possibly) unclear on how to apply it. In these disputes, the ECJ ruling is effectively internalized by the national court and enforced through domestic mechanisms. The distinction between international ruling and domestic compliance is blurred.

Infringement disputes span many topics, especially those which speak to core aspects of European integration. For example, these disputes have addressed the

free movement of goods, freedom of movement of persons, provision of services, equal treatment and social rights, etc. Notably, many of the disputes have concerned economic integration and conflicts over trade-related aspects of the internal market. As a result, the ECJ is thought to play a key role in sustaining European economic cooperation (Börzel, 2001; Panke, 2007).

The typical infringement dispute concerns a European Commission directive, an order for member countries to achieve a policy objective through their own national legislation or rules. Member states have some leeway in how they implement or “transpose” the directive, but they have a clear obligation to do so promptly, correctly, and to report their implementation to the European Commission. The Commission monitors implementation and inquires about potential failures. If the member state does not provide evidence that it has implemented the directive, the Commission can then bring formal legal charges at the ECJ.¹¹ A full lawsuit follows where a panel of judges evaluate the case on the merits and issue a verdict or “reasoned opinion” (EC Article 169). The ECJ typically issues its judgment within two years of the initial referral to the court.

At this point, if the ECJ determines the defendant breached EU law, that government is obligated to correct its laws or practices and resolve the initial dispute as soon as possible. Continued noncompliance prompts a second round of litigation. The Commission can ask the Court to impose penalties which are retroactive and may apply for the full duration of the violation. Again, the ECJ evaluates the case on the merits and issues a judgment (EC Article 228). Additional financial sanctions may be imposed in the case of continued noncompliance.¹²

The European Union requires prompt compliance with these judgments and uses penalties to induce cooperation. The ECJ website explains: “if the country is found

¹¹This type of lawsuit can also be started by another member country but this is rare.

¹²Article 260 of the Treaty on the Functioning of the European Union and Article 106a of the Euratom Treaty.

to be at fault, it must put things right at once. If the Court finds that the country has not followed its ruling, it can issue a fine” that is large enough to “ensure the penalty itself has a deterrent effect.”¹³ Despite the emphasis on timely compliance, in some instances disputes stretch on for years. For example, when France was sued over its ban on foreign health-care worker in French overseas territories and lost, it resisted the Court’s judgment for nearly nine years.¹⁴ In another instance, the Commission sued Greece over its failure to adopt EU policy on professional certifications.¹⁵ Greece did not comply and was sued again under noncompliance procedures. This conflict lasted more than seven years.

4.2.4 Comparing Institutional Design

The design of international institutions is widely thought to shape countries’ decisions to comply with their international commitments (e.g. Abbott and Snidal 1998; Koremenos, Lipson, and Snidal 2001; Carrubba 2005; Gilligan, Johns, and Rosendorff 2010). International courts, in particular, vary in their design with some more flexible than others. These institutions are flexible if they allow countries to sometimes violate their treaty commitments or provide latitude to negotiate, delay compliance, and compensate one another for temporary transgressions (Rosendorff and Milner, 2001). By contrast, institutions are rigid if they restrict negotiation and compensation, or allow countries little control over the timing of disputes and impose penalties for noncompliance.

The WTO dispute settlement mechanism is a relatively flexible institution in several respects. First, it provides many legal mechanisms by which countries can delay action and prolong their disputes. Defendants can demand an extended grace

¹³See: *Europa* and *The European Commission at Work*.

¹⁴Case C-263/88 *Commission v. France*.

¹⁵Case C-365/93 *Commission v. Hellenic Republic*.

period (RPT), request arbitration, appeal rulings, or take other action to extend the duration of the dispute. This gives defendant countries ample room to extend the duration of litigation and, even after a ruling has been issued, delay the adoption of new policies needed to implement the ruling.

Second, the WTO does not issue retrospective punishments or issue injunctions. That means that as long as the dispute is active, the losing defendant can keep the contested policy in place, often with great benefits to the sector or industry that enjoys the illegal trade protection. Some scholars argue this creates a “remedy gap” in the WTO that gives defendant governments the option to violate trade rules for several years without facing retaliation or other penalties (Brewster, 2011). One consequence is that defendants can—and often do—use time to their advantage, prolonging legal proceedings and increasing the plaintiff’s cost.

Third, the WTO dispute settlement mechanism is flexible because enforcement is decentralized. Before any adverse ruling can be enforced, the plaintiff government, not the WTO itself, must seek additional legal recourse. The plaintiff can demand that the defendant provides compensation, pending full implementation, in the form of a tariff reduction or other trade benefit that offsets the harm from the defendant’s continued violation.¹⁶ Or it can request authorization to retaliate against the defendant (suspension of concessions) in the form of trade sanctions meant to induce compliance. Disputants may negotiate settlements and rulings are only enforceable insofar as the plaintiff is willing or able to do so. The burden is on the plaintiff to generate economic incentives for the defendant to comply so plaintiffs with little leverage (i.e. small economies) will inevitably find the process more burdensome and less successful. Decentralized enforcement contributes to the WTO’s flexibility by allowing some defendants to “get out of jail free” when they are willing to hold out long enough.

¹⁶WTO Dispute Settlement Understanding, Article 22.2.

This is compounded by the WTO's lenient monitoring practices. Defendants that lose WTO disputes are obligated to periodically report on their progress toward compliance, but the institution does not independently verify that the government has followed through. Again, the WTO relies on the plaintiff to pursue the case further if compliance is inadequate. This gives the defendant more opportunity to evade enforcement and retain non-compliant policies.

By contrast, the European Court of Justice has a more rigid design. First, the ECJ has strict time-lines for its legal procedures, allowing defendants few opportunities to delay.¹⁷ Defendants cannot unilaterally stay legal proceedings¹⁸ and, once the oral arguments stage of the dispute has begun, they have limited ability to postpone hearings to produce evidence or otherwise slow the procedures.¹⁹ Opportunities for appeal are limited. Moreover, ECJ rulings are binding and immediately applicable,²⁰ offering defendants no grace period for implementation as is standard under the WTO.

Second, the ECJ can impose retrospective punishments on the defendant, reducing the benefits of delay. In some instances, the European Commission has threatened obstinate governments with daily fines going back for the entire duration of the violation. The ECJ can also order the defendant to cover certain costs associated with litigation, as an additional penalty.²¹ As compared to the WTO, where timing can be an avenue for maneuverability, the ECJ restricts the defendant's ability and incentive to prolong legal disputes.

¹⁷European Court of Justice Rules of Procedure, Articles 49 through 52.

¹⁸Ibid, Article 55.

¹⁹Ibid, Article 135.

²⁰Ibid, Article 91.

²¹European Court of Justice Rules of Procedure, Chapter 6.

Third, these restrictions are possible because the European Union has a centralized approach to enforcement. Its administrative branch, the European Commission, has bureaucratic autonomy, and offers countries little opportunity to negotiate informally (Carrubba and Gabel, 2015). The Commission monitors how governments correct their violations in response to adverse ECJ rulings. In the case of prolonged noncompliance, the Commission can request authorization from the ECJ to penalize the defendant government with financial sanctions. When calculating the amount, the Commission accounts for the wider effect of compliance on the European community and penalties are supposed to be large enough to have a deterrent effect. This centralized approach to enforcement does not rely on the motivation or ability of an individual plaintiff government to seek penalties. The result is that enforcement is regimented and presents a credible threat to defendant governments that are tempted to defy the court.

The differences between these institutions have domestic implications. Under the WTO, governments have greater opportunity to deal with domestic political and economic circumstances. When domestic political pressure for noncompliance is short-lived—for example, if voters in a key district for an upcoming election prefer trade protection—politicians can use the WTO’s remedy gap to their advantage. They may delay compliance until that domestic pressure has passed. Under the ECJ, governments have less opportunity to deal with domestic opposition. They cannot easily delay compliance without incurring substantial penalties. These implications are developed below.

4.3 Institutional Design and the Domestic Politics of International Cooperation

4.3.1 Domestic Politics of International Cooperation

Domestic groups, especially industries, have preferences over international economic cooperation. International agreements constrain the set of policies that member governments can implement, including policies that favor certain domestic groups. The WTO restricts members' trade policies by placing limits on tariffs, subsidies and other non-tariff barriers. This limits a member's ability to respond to domestic demands for trade protection. Import-competing domestic industries may demand trade protections that stave off foreign competition but WTO rules tie a government's hands, limiting the instances in which a government can respond to these demands. Similarly, the European Union eliminates standard trade barriers and constrains members' policies in many other economic domains including banking, competition policy, enterprise, taxation, and other matters that have trade implications. For example, domestic firms may gain a competitive advantage from tax schemes that drive up the price of imported products from other EU members and will have an incentive to oppose EU rules. In each instance, domestic groups have well-formed preferences over international cooperation, depending on how those rules affect them.

Countries vary in the degree to which political authority is concentrated within government. When domestic political institutions have many checks and balances, authority is less concentrated and more political actors can block policy change. Division in authority can also arise from party politics. When the domestic government features a relatively strong opposition party or many opposition parties, there are significant obstacles to policy change. Together, domestic institutional and partisan divisions form *veto points* in government. As veto points increase, political authority becomes more fragmented and policy becomes more difficult to change (Tsebelis, 1995; Henisz, 2000; Tsebelis, 2003; Henisz and Mansfield, 2006).

Veto points determine how responsive governments are to the demands of different political actors. When the government has many veto points, policy change is quite difficult because it requires coordination among many political actors. In these instances, it is more likely that any one political actor can block the adoption of a new policy. For example, when a proposed trade policy hurts them, industries may voice their opposition and pressure politicians to obstruct its implementation. Conversely, when an existing trade policy benefits them, industry groups may urge politicians to block changes that reduce or eliminate those benefits. The more veto points in government, the more likely they are to succeed.

These domestic constraints can affect international cooperation at several stages. Multiple veto points can make it more difficult for countries to join international agreements (Milner and Rosendorff, 1996, 1997; Simmons, 2009). They narrow the set of international agreements that can be ratified because when a government attempts to sign an international agreement, legislators or other political actors can obstruct its domestic adoption. Conditional on joining an agreement, domestic veto points can make it less likely that a country will violate its international commitments. For instance, sometimes domestic interest groups demand trade protection policies that violate international agreements. A government with substantial divisions in authority may be less responsive to such demands because the groups will have to persuade many political actors to support a new policy (Rickard, 2010). These tendencies imply that countries with many domestic veto points are apt to make more credible international commitments about future cooperation (Mansfield, Milner, and Rosendorff, 2002). Accordingly, many scholars have concluded that domestic veto points promote international cooperation.

However, countries can and do violate their international commitments, particularly when domestic pressure to do so is acute (Johns and Rosendorff, 2009). The prevalence of WTO and ECJ lawsuits supports this point. When a country violates its international commitments and loses a lawsuit at an international court, it is

obligated to correct the violation. Domestic institutional and partisan constraints can block the adoption of policies that will bring a country into compliance. Conditional on an initial violation, domestic veto points should make it more difficult for a government to implement an adverse ruling from an international court and return to compliance.

International disputes do not end with a single instance of noncompliance. When a country defies the ruling of an international court, it will encounter many opportunities to comply and has an obligation to follow through as long as the violation remains in place. Long-lasting disputes are usually instances where the defendant resisted the court ruling and the parties went through multiple rounds of appeals and compliance proceedings. Defendant governments ordered to change their policies may “drag their feet” and delay. At the same time, domestic constraints apply over time: so long as a leader faces many domestic veto players, she should continue to find it difficult to reverse the violation. Noncompliance will tend to persist until the political or economic conditions change and the demand for the impermissible policy subsides. As a result, defendant governments with many veto points should tend to have longer-lasting disputes.

4.3.2 Relationship Between Domestic Politics and Institutional Design

Defendant governments that lose international lawsuits are obligated to comply, even when they face substantial domestic hurdles. International courts can authorize a plaintiff to penalize the defendant, creating international pressure to comply. International courts that provide greater flexibility allow losing defendant governments to delay compliance, partially correct violations, and/or provide temporary compensatory measures. These courts supply greater opportunity for defendant governments to respond to domestic pressure and less incentive to overcome the domestic hurdles. By contrast, international courts that are more rigid make few such allowances.

These courts provide little opportunity for losing defendant governments to delay and manage domestic political constraints.

When a defendant government has more opportunity to resist compliance and delay, thanks to a flexible institution, the effect of domestic veto players will be more pronounced. In these cases, some defendant governments will use institutional flexibility to its full extent, while domestic veto players obstruct policy reforms.²² The extent of domestic constraints should have an important impact on dispute resolution, with more constraints reducing the likelihood of compliance at any point in time. The rates of compliance for more constrained governments should differ substantially from compliance rates for less constrained governments. Conversely, under rigid institutions, compliance rates for more and less constrained governments should differ little. Over time, these same tendencies should show up as *dispute duration*, defined as the litigation period plus, in the event of an adverse ruling, the period of noncompliance. Defendant governments that face many domestic constraints should have longer-lasting disputes than less-constrained governments do, and this discrepancy should be more pronounced under flexible institutions. All else equal, it follows that:

1. A defendant government with more domestic constraints will engage in longer lasting disputes.
2. The larger the international punishment, the smaller the impact of domestic constraints on dispute duration.

Together, these hypotheses predict that when international penalties are small and domestic politics is divisive, litigation will be long and drawn-out. And instead

²²For example, when Canada lost a WTO ruling about the magazine industry, the Canadian government appealed to buy time and stall domestic political backlash. One newspaper said, “the four- to five-month delay [from appeal] could provide political advantage for Canadian officials. ‘The main purpose of the appeal is to get the darn thing out of the way until after the election,’ said Gordon Ritchie, a chief Canadian negotiator of the U.S.-Canada Free Trade Agreement.” See: “Canada to ignore magazine decision” *Journal of Commerce*, March 16, 1997.

of complying immediately, a defendant that faces many veto points will delay until domestic costs subside. Domestic divisions obstruct compliance with an international court's ruling and this effect is most prominent when the penalties for noncompliance are small.

4.4 Empirical Analysis

4.4.1 Data Sources

To evaluate the link between domestic politics and international legal disputes, I create two datasets. The first is a dataset on all legal disputes filed at the WTO between 1995 and 2012.²³ Drawing on data from Horn and Mavroidis (2008), I identify the plaintiff and defendant countries, the legal stages in each dispute, start and end dates, and legal rulings. There are 404 disputes in this period, covering a wide range of trade topics, products and engaging more than 40 countries. I include information on the total bilateral trade between the plaintiff and defendant as well as the defendant country's gross domestic product in the year the dispute was initiated (United Nations, 2013; World Bank, 2013).

My explanatory variable is domestic veto players in the respondent government that arise from institutional checks and partisan opposition. I measure this as VETO POINTS using the Political Constraints Index (Henisz, 2002). It accounts for the number of independent branches of government, the extent of partisan alignment across branches of government, and preference heterogeneity within each legislative body. When measuring independent branches of government, it accounts for federal institutions. When measuring partisan alignment, it accounts for party composition and left/right preference which change over time. VETO POINTS range from zero (least constrained) to one (most constrained). This metric is ideal because

²³I truncate the sample at 2012 to allow sufficient time to observe dispute resolution.

it has comprehensive coverage across countries and over years in my sample and is widely-accepted among political scientists. It allows me to draw comparisons between otherwise dissimilar countries, providing a unified summary of domestic obstacles.

The second dataset identifies infringement disputes filed at the European Court of Justice from 1978 to 1999. I select the subset of ECJ infringement disputes that concern trade policy and topics relating to European commerce such as competition policy, taxation, etc. in order to facilitate comparisons between the two institutions. I draw on data collected by Börzel and Knoll (2012) and coded according to legal stages, issue area, start and end dates, and legal rulings.²⁴ There are 673 such disputes involving the 15 core EU member countries.²⁵ I include the same VETO POINTS variable as above. I also include information on the defendant countries' trade dependence and gross domestic product in the year the dispute was initiated (World Bank, 2013).

4.4.2 Disputes Stages and Veto Players

Table 4.1 presents a summary of the legal disputes at the World Trade Organization Dispute Settlement Mechanism and at the European Court of Justice, grouped by stage of legal proceeding. Both institutions manage to resolve a large share of disputes early, without a legal verdict. At the WTO, approximately 43% of cases go on to receive a ruling whereas at the ECJ 53% do. There are many reasons a dispute might be resolved early. First, it is possible that the allegations were unsubstantial; the defendant government had not, in fact, violated its obligations. In other instances, the threat of legal recourse compels the defendant to correct a vi-

²⁴Börzel and Knoll (2012) have not yet released data after 1999 and so to ensure consistent coding, I end my sample at 1999 as well. I select infringement disputes that dealt with EU directives over agriculture, budgets, competition, enterprise and industry, internal market, taxation and the customs union.

²⁵These countries are Austria, Belgium, Germany, Denmark, Spain, Finland, France, UK, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal and Sweden.

olation, bringing its policies and practices into compliance. The plaintiff may then drop the charges. In yet other instances, the disputants may settle without further legal intervention.

When the countries complete litigation, the international court will issue verdicts that establish fault and specify terms of compliance. In both courts, the vast majority of rulings favor the plaintiff on at least one legal claim: 90% and 88% of WTO disputes and ECJ trade infringement cases, respectively. This means that in almost all instances, defendants are required to change their policies in response to the international court's ruling. This may entail removing trade barriers that violate WTO obligations or adopting new policies that follow the mandates of EU law on economic integration.

Table 4.1: Disputes at WTO, 1995-2012 and ECJ Infringement Disputes, 1978-1999, by Stage of Legal Proceeding

LEGAL CATEGORY	DISPUTES (#)		VETO POINTS <i>mean (sd)</i>	
	WTO	ECJ	WTO	ECJ
<i>First Stage</i>				
Total Disputes	404	673	0.41 (0.14)	0.49 (0.13)
Disputes with Rulings	170	359	0.40 (0.15)	0.48 (0.13)
Adverse Rulings	153	315	0.40 (0.14)	0.48 (0.13)
<i>Second Stage</i>				
Non-Compliance Referrals	26	17	0.45 (0.08)	0.43 (0.10)
Adverse Non-Compliance Rulings	9	4	0.45 (0.08)	0.37 (0.09)
<i>Second Adverse Ruling Rate</i>	<i>5.9%</i>	<i>1.3%</i>	–	–

Note: WTO counts include all disputes. ECJ counts are restricted to infringement disputes on trade-related topics that were referred to the court for a ruling. Veto points are for the defendant government at the beginning of dispute.

Many legal disputes end at this point, when the defendant government implements the ruling within the established time span. Occasionally, the plaintiff pur-

sues the case further when it believes the defendant government has failed to take adequate measures. In these instances, defendants are accused of failing to comply with the legal rulings (second-order noncompliance). Both institutions have legal procedures for dealing with allegations of continued violations. Noncompliance referrals are more common at the WTO (17% of cases) than the ECJ (5.4% of cases).²⁶ Usually, these cases are resolved after the referral. The added pressure of a continued lawsuit may prompt some defendants to comply or expedite a settlement agreement. Some disputes, however, go on to receive a second legal verdict of continued noncompliance and at this point the defendants are subjected to penalties. This has happened in 9 cases at WTO, comprising 5.9% of the initial adverse rulings, and only 4 cases at the ECJ, comprising 1.3% of the initial adverse rulings. The legal process translates into long-lasting disputes where both parties stand their ground.

Table 4.1 also displays average domestic veto points of defendant governments in each stage of legal disputes. Defendant countries that clearly defy the WTO rulings tend to have slightly higher domestic veto points than the typical WTO defendant. At the ECJ, the pattern is reversed but none of the differences are statistically significant at conventional levels. In both institutions, countries with many domestic veto points are just as likely as their less constrained counterparts to receive an adverse ruling.²⁷

On first glance, one might think that the rarity of disputes with a second adverse ruling is good news about compliance. However, this almost surely understates the rate at which countries ignore rulings and continue to violate their treaty obligations. There are several reasons. Sometimes plaintiffs may drop the lawsuit because

²⁶This discrepancy may be explained by higher compliance rates at the ECJ than the WTO. It may also be explained by the European Commission's relative reluctance to sue a member government a second time, once that government has already defied the ECJ. Additional work is needed to differentiate between these mechanisms.

²⁷The exception is for the ECJ's second noncompliance disputes. Countries with few domestic veto points are more likely to experience adverse rulings at this second stage.

pursuing the complaints against a recalcitrant defendant is costly and futile. Other times defendants may falsely claim to have complied. Uncovering evidence of continued violations may be prohibitively costly for the plaintiff and again, she may regard it preferable to drop the lawsuit. Either way, noncompliance may be common and go unchallenged through the standard legal channels. This may lead to bias *against* finding a veto players effect (i.e. biased toward a null effect), because even governments with many veto points that defy the court will be presumed compliant. One way to gain insight into defendant government's response to adverse rulings is to examine dispute duration, or how long a formal legal dispute lasts before being satisfactorily resolved.

4.4.3 Dispute Duration and Veto Players

Measured from the time a plaintiff files its initial lawsuit to the time the case is resolved, WTO disputes last longer than ECJ disputes. This time period can be divided into two stages. The first is the dispute period, starting with the plaintiff's formal legal complaint and ending with the court's ruling. Some cases are resolved during the first period, without a ruling. Sometimes the ruling favors the defendant and the dispute ends. Second, when the court rules against the defendant and determines that government has violated its treaty commitments, there is a noncompliance period, when the defendant government must implement the ruling through domestic policy. In the latter period, as long as the dispute remains unresolved, I assume the initial violation remains in place. This is a reasonable assumption because when the defendant satisfactorily complies, it reports the action to the plaintiff or international court and thereby nullifies international penalties.

WTO disputes with a ruling typically last 3 years with some lasting as many as 15 years. At the ECJ, the average infringement dispute with a ruling is resolved in 2.4 years with some lasting 8 years. For disputes with allegations of continued

noncompliance (a second round of litigation), the average dispute duration is longer at 4.8 and 3.5 years, respectively.

To model dispute duration under each institution, I run a survival analysis using Cox proportional hazard models. The models evaluate the relationship between domestic veto points in the defendant government and the “risk” of dispute resolution. The dependent variable is the dispute survival from the initial lawsuit to resolution and the explanatory variable is the defendant government’s domestic veto points.

For the WTO, I control for the defendant’s gross domestic product (`DEFENDANT GDP`) in logged units and trade dependence (`DEFENDANT TRADE DEPENDENCE`). I expect that defendants with larger economies are likely to engage in longer disputes—less likely to capitulate—because they are relatively less vulnerable to retaliation from the plaintiff.²⁸ To reflect the bilateral nature of WTO disputes, I also control for the total trade between the plaintiff and defendant countries (`BILATERAL TRADE`). In some specifications, I include an indicator where the European Union is the defendant (`DEFENDANT IS EU`) and control for the plaintiff’s GDP in logged units (`PLAINTIFF GDP`).

For the ECJ, I control for the defendant country’s `GDP`, `GDP PER CAPITA`, and `TRADE DEPENDENCE`.²⁹ In one specification, I use country fixed effects to isolated within-country variation, since each country’s veto points fluctuate (to a small extent) over time. The results are displayed in Tables 4.2 and 4.3.

²⁸My expectations about trade dependence are less clear. On one hand, highly trade dependent countries may be more determined to retain disputed trade policies and less willing to capitulate. On the other hand, these same countries may be more vulnerable to threats of enforcement by the plaintiff, since even a minor suspension of concessions can substantially harm a trade-dependent country.

²⁹Some literature on the European Court of Justice suggests that compliance issues are largely a “Southern problem,” meaning that the less wealthy Mediterranean countries are the least likely to follow through on their EU obligations and comply with ECJ rulings. By this account, GDP per capita should be negatively associated with dispute resolution.

Table 4.2: Cox Proportional Hazard Model of WTO Disputes, 1995-2012, where Event is Dispute Resolution

	Dispute Duration		
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>
Defendant Veto Points	−2.400*** (0.441)	−2.306*** (0.469)	−2.003*** (0.484)
Defendant GDP (log)		0.133*** (0.044)	0.169*** (0.047)
Defendant Trade Dependence		0.003 (0.003)	0.008** (0.003)
Defendant is EU			−0.524** (0.213)
Bilateral Trade [†]		−0.088 (0.114)	−0.086 (0.116)
Plaintiff GDP (log)			0.015 (0.036)
Observations	401	401	398
Log Likelihood	−1,232	−1,226	−1,216

Note: Cox-proportional hazard estimates were calculated with the **Survival** package in R. Dispute duration was calculated from the initial complaint to the date of settlement, compliance, or, in the case of unresolved disputes, January 1, 2012. [†] denotes normalized variable. Significance codes *p<0.1; **p<0.05; ***p<0.01

Table 4.2 presents estimates from survival models of WTO disputes where the “event” is dispute resolution. All estimates are based on Cox proportionate hazard models where the dispute begins with the date of the lawsuit³⁰ and ends with the date of resolution through settlement, compliance, or for unresolved disputes, the censor date.³¹ The analysis shows that defendant governments with many domestic veto points engage in longer-lasting disputes, as demonstrated by the negative coefficient. The European Union tends to engage in longer-lasting disputes but on average, defendants with larger economies seem to resolve their disputes more quickly, either through settlement or compliance. The total trade between the plaintiff and defendant and the plaintiff’s GDP are not significant predictors of dispute duration.³²

Table 4.3 presents estimates from survival models of ECJ disputes using a similar approach. Domestic veto points are negatively associated with dispute resolution but the correlation is only significant when control variables and country fixed effects are introduced. The country-fixed effects control for some of the cross-country variation, allowing me to isolate the smaller within-country fluctuations that correlate with ECJ infringement rulings and domestic veto points. The defendant country’s trade dependence is consistently associated with prompt dispute resolution. The defendant’s GDP and per capita GDP are not consistently associated with dispute resolution in the survival models.³³

³⁰Specifically, at the WTO this is the date of the request for consultations and at the ECJ this is the date of referral to the court.

³¹For both institutions, some disputes remained unresolved and therefore lacked a clear end date. In these cases, I censored the duration. For the WTO, the censor date is January 1, 2012. The results are robust to other choices of censor dates. For the ECJ cases, the duration is censored as the last date of data collection by Börzel and Knoll (2012).

³²Results are similar when I restrict the analysis to the post-ruling duration, measured from the date the panel delivered its judgment to the end date of the dispute.

³³Results are similar when I restrict the analysis to the litigation duration or the post-litigation duration, measured from the date the ECJ issues a judgment to the end date of the dispute.

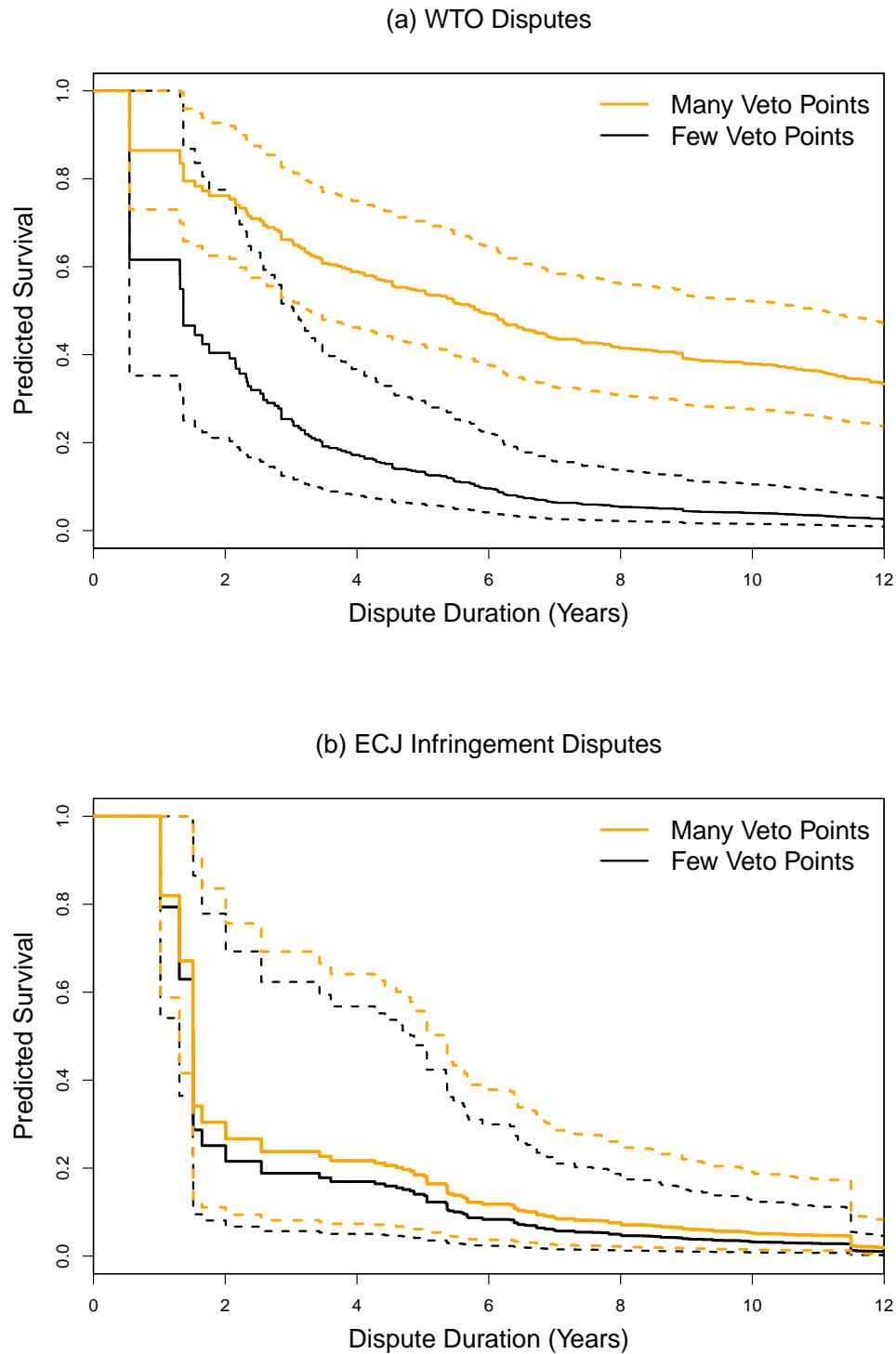
Table 4.3: Cox Proportional Hazard Model of ECJ Disputes, 1978-1999, where the Event is Dispute Resolution

	Dispute Duration		
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>
Defendant Veto Points	0.483 (0.325)	−0.250 (0.429)	−1.221* (0.740)
Defendant GDP (log)		0.201** (0.100)	−4.631 (3.172)
Defendant GDP per Capita (log)		−1.177*** (0.388)	−1.392 (3.483)
Defendant Trade Dependence		0.010*** (0.003)	0.030*** (0.009)
Country Fixed Effects?	No	No	Yes
Observations	671	671	671
Log Likelihood	−2,761	−2,753	−2,720

Note: Cox-proportional hazard estimates are calculated with the **Survival** package in R. Dispute duration is calculated from the initial complaint to the date of compliance, or in the case of unresolved disputes, the latest date of data collection. All controls are for the defendant country. Significance codes *p<0.1;**p<0.05; ***p<0.01.

To facilitate comparison between the two institutions, Figure 4.1 illustrates predicted dispute duration based on the survival models above. Predictions are calculated using many and few veto points with control variables held at their means. Specifically, “many” veto points is set at the fourth quantile of the distribution of observations and “few” veto points is set at the first quantile. The plots show 95% confidence intervals around the estimates, denoted by dashed lines.

Figure 4.1: Predicted Survival of Legal Disputes with Domestic Veto Points



Note: Survival curves were fit using Cox proportional hazard function in **Survival** package in R and model specifications included all controls held at their means. Dotted lines denote 95% confidence intervals.

Dispute duration varies with the defendant government's domestic political constraints. In both institutions, defendant governments with more domestic veto points are predicted to engage in longer-lasting disputes. The relationship between domestic veto points and dispute duration is more pronounced for WTO dispute than for ECJ disputes. When the defendant government has many veto points, WTO disputes tend to last significantly longer. This is shown by the significant separation of survival curves in Figure 4.1(a). At the ECJ, having more domestic constraints is weakly associated with longer-lasting disputes but the association is not statistically significant and barely discernible from Figure 4.1(b).

4.4.4 Illustrations

Two examples, one from the WTO and one from the ECJ, illustrate the impact of domestic constraints and institutional design on international legal disputes. To improve comparability, I chose disputes with closely related topics, overlapping disputants, and similar time frames. In both cases, the courts ruled against the defendants, mandating removal of a barrier to trade for a specific product (meat) that was ostensibly in place to protect consumers. In both instances, the defendant initially resisted the court ruling and encountered multiple veto points. The WTO dispute was drawn out and remains largely unresolved; the ECJ dispute was dealt with relatively quickly, despite concerns that there was strong domestic opposition to complying.

WTO Dispute: US and Canada v. European Union – Hormone-Treated Meat

The first example is a 1996 WTO dispute where Canada and the United States sued the European Union over its ban on hormone-treated meats.³⁴ The WTO ruled against the EU, ordering it to comply by May 1999. Within Europe, there

³⁴DS 26 and 48: EC – Hormones, 1996

was staunch resistance to the ruling. One British newspaper said of the controversial ruling, “political leaders are worried about provoking a consumer backlash [by complying]—and upsetting the powerful farming lobby.”³⁵ Domestic politics formed a major obstacle which was exacerbated by institutional barriers within the European Union’s bureaucracy. Implementing the ruling would require that the European Council adopt an amendment to the EU Hormones Directive, the source of the initial dispute. Because the European Council is composed of heads of state of the (then) fifteen EU member governments and leadership from the European Commission, and decisions mostly take place through consensus, there are many opportunities for veto players to block compliance. Even when the Council takes action, member governments, each of which had a political stake in the dispute, might resist implementing the decision correctly at the domestic level. With many veto points, the EU failed to implement the ruling and defied the WTO.

The relative flexibility of the WTO did not help persuade the EU to comply. The United States and Canada sought retaliation and won \$116.8 million and \$11.3 million in remedies, respectively, in July 1999. At this point the European Union ceased providing status reports on its implementation, apparently in protestation, making it difficult for the WTO to track implementation.³⁶ The deadlock dragged on for years. By 2003, the European Council amended the Hormones Directive and claimed to be in full compliance with the adverse ruling, leading to another round of ineffectual litigation.³⁷ One Canadian newspaper noted, “it is unlikely that the latest

³⁵ “Europe Faces Sanctions over Hormone Beef Ban,” *Birmingham Post*, United Kingdom, June 4, 1999.

³⁶ WTO Dispute Settlement Body Meeting Minutes WT/DSB/M/282, p.8.

³⁷ See Directive 2003/74/EC which entered into force 14 October 2003. In November 2003, the EU requested Canada cease retaliation. Canada refused and the EU brought a counterclaim, intending to show that Canada’s continued retaliation was unwarranted.

ruling will persuade the European countries to drop their bans.”³⁸ After more than a decade of impasse, the disputants signed bilateral agreements and finally settled in 2011.³⁹ According to the WTO meeting minutes: “the United States had long stressed the importance of flexibility in the WTO dispute settlement system. The ability of the United States and the EC to conclude this [agreement] was a demonstration of how that flexibility could be put to use by members to make progress in a long-running dispute”⁴⁰ Despite this positive assessment, the flexibility also allowed for one of the more domestically fraught political processes the WTO has seen. It allowed the many veto points in the EU to lock the violation in place and the EU did not comply, substituting instead the equivalent of a cease-fire agreement.⁴¹

ECJ Dispute: European Commission v. France – Beef Imports

In the second example, a dispute at the European Court of Justice, the defendant government faced domestic obstacles to compliance and defied the ruling for a short time before capitulating. The initial problem arose when, in response to the mad cow disease crisis, European countries placed an embargo on British beef exports. While the rest of the EU lifted the ban in 1999 after the immediate threat had passed, the French government refused. The European Commission, prodded by the UK, sued France at the ECJ, which ruled that France’s refusal to lift the embargo was illegal and a clear breach of EU law.⁴² France ignored the ECJ, inciting a heated political

³⁸“Canada Wins WTO Ruling over European Ban of Hormone-Treated Beef Imports,” *Prince George Citizen*, British Columbia, Canada, March 29, 2008.

³⁹Memorandum of Understanding between the US and EU, signed March 2009 and between Canada and the EU, signed March 2011.

⁴⁰WT/DSB/M/270, June 19, 2009, p.18.

⁴¹Note that the disputants have renegotiated settlement terms as recently as 2014, suggesting the 2011 solution was not decisive.

⁴²Commission v. France, Case C-1/00, Judgment of the Court December 13, 2001.

debate. As the Guardian reported, “France must now comply with the ruling...or face a fine from the European Commission on each day that it fails to act, which could run into hundreds of thousands of pounds...[a likely scenario] in light of the country’s powerful farming lobby and impending presidential and parliamentary elections.”⁴³ With staunch domestic opposition and high political stakes, the French government attempted to evade the ruling.⁴⁴ And the numerous institutional constraints in French government made compliance particularly difficult.⁴⁵

However, the relatively rigid design of the ECJ gave France little chance to flout European law without significant penalty. The European Commission sought a daily fine equal to £100,000. As one newspaper noted, “Within weeks of the fine being urged, France’s food standards agency recommended the ban was removed” and the French government finally lifted its trade barriers.⁴⁶ The EU enforcement measures were sufficient that the French government overcame its domestic political constraints and fully complied before the year ended.⁴⁷

The examples demonstrate that when institutions are relatively flexible, imposing a small punishment on governments that defy its rulings, domestic constraints can lead to a prolonged disputes. When institutions are relatively rigid, penalizing

⁴³ “French Ban on UK Beef Ruled Illegal,” *The Guardian*, December 14, 2001.

⁴⁴ “French find new ways to keep out British beef,” *Sunday Suteleaph*, London, October 20, 2002.

⁴⁵ France has a semi-presidential system with both upper and lower houses of parliament. It had a moderate number of veto points at this time, 0.44. This ECJ dispute came in the months leading up to the 2002 elections.

⁴⁶ “France Finally Lifts Its Illegal Ban on Imports of British Beef,” *Aberdeen Press*, October 26, 2002.

⁴⁷ The Commission’s 2002 annual report states: “As France did not take measures to give effect to the judgment given against it by the Court of Justice on 13.12.2001 in Case C-1/00, for refusing to lift the embargo on imports of British beef properly marked or labeled from 30 December 1999, the Commission commenced expedited Article 228 proceedings on 21.3.2002. The infringement was put right after the Commission applied to the Court to give judgment against France for failure to comply with the earlier judgment, with a periodic penalty payment in support. The Commission withdrew its action once the infringement was terminated.” See: 20th Annual Report on Monitoring the Application of Community Law, November 21, 2003.

defendant governments that defy its rulings, a government is more likely to overcome domestic constraints and comply.

4.5 Discussion

Domestic political constraints and international obligations can pull governments in different directions, particularly when they find themselves embroiled in international legal disputes. Domestic politics should affect the process of international dispute resolution by constraining leaders and making it more difficult for a defendant government to comply with an adverse legal ruling.

This tendency is evident in the dispute records at two of the most active and prominent international judicial bodies, the World Trade Organization's dispute settlement mechanism and the European Court of Justice. Although domestic veto players in the defendant government are not consistently associated with more legal stages in the disputes, including noncompliance proceedings, they are associated with longer-lasting disputes. The total dispute duration, including the litigation period and, in the case of adverse rulings, the noncompliance period, is longer when defendant governments have many domestic constraints. In WTO lawsuits, defendant governments with many domestic institutional checks and strong partisan opposition tend to engage in disputes that drag on for many years and often remain unresolved. Infringement disputes at the European Court of Justice display similar but weaker patterns. When defendant governments lose international disputes and are urged to change their domestic policies, veto players appear to obstruct the policy reform process, prolonging noncompliance.

When international courts are more flexible, allowing defendant governments greater leeway to accommodate domestic political pressures and constraints, defendants take advantage of these opportunities. The ill-effects of domestic constraints on international dispute resolution is more pronounced under the WTO than it is

under the ECJ. I argue that this difference is due, at least in part, to the design of the institutions. The WTO Dispute Settlement Mechanism is relatively flexible, allowing defendant governments significant opportunity to delay compliance and circumvent adverse legal rulings. With relatively weak enforcement mechanisms, defendant governments that face many veto players have both the opportunity and incentives to delay and defy the court. In this context, domestic veto points are strongly associated with dispute outcomes. The European Court of Justice is relatively rigid, permitting losing defendants little opportunity or incentive to defy adverse rulings. The institutional design appears to homogenize dispute outcomes, and domestic veto points appear to play only a modest role.

Flexible institutions are thought to have many advantages, including stabilizing and promoting long-term cooperation and, in some instances, permitting deeper forms of cooperation. While the literature on the optimal design of international institutions focuses largely on trade-offs between features such as depth, flexibility, and stability, (Rosendorff, 2005; Johns, 2014, 2015), less attention has been paid to the link between design features and domestic politics. The evidence above highlights an important disadvantage of flexible institutions. International courts that allow countries to use flexibility to accommodate domestic pressure may also pave the way for opportunistic behavior. When a defendant government is faced with an adverse ruling, it may extend the duration of the dispute and avoid compliance as long as it can. And in some cases, this means a defendant government that is beholden to domestic groups that oppose the ruling may fail to overcome these hurdles and remain non-compliant. Because opportunistic behavior can sometimes create substantial costs for plaintiffs or other countries, it risks undermining the cooperative behavior that defines these international economic organizations. In these instances, ample institutional flexibility may have real negative consequences for international cooperation.

CHAPTER 5

Conclusion

5.1 Domestic Politics and the Effectiveness of International Courts

This dissertation started with the question, when do international courts work? The question is at the center of a long-standing debate that is no less divisive today than it was decades ago when international organizations were first proliferating. Some scholars argue that international institutions are merely expressions of underlying power relations among countries—expressions of state preferences but little else. Ultimately ineffective, they offer “false promises” about international cooperation (Mearsheimer, 1994). According to this view, international courts work only insofar as they articulate in legal terms the behaviors that the more powerful countries would have demanded anyway.

Other scholars argue that international institutions carry significant normative authority and prompt real changes in state behavior (Finnemore, 1993; Finnemore and Sikkink, 1998). The judgments of courts are binding legal requirements which guide countries’ conceptions of acceptable and unacceptable international behavior.

Still other scholars take the position that these institutions are rational solutions to the coordination and collaboration problems that plague international affairs.

They have the capacity to shape state behavior by changing incentives.¹ Courts can work when they provide information to governments, deescalate conflict, or facilitate enforcement of rules (Busch and Reinhardt, 2000; Carrubba, 2005; Johns, 2011). They highlight countries' violations of primary obligations, issue rulings that specify necessary reforms, and authorize sanctions that make noncompliance costly. This project draws heavily on the latter perspective, considering the conditions under which international courts incentivize government to cooperate.

Like previous studies, this dissertation confronts an important methodological dilemma in studying the effectiveness of international institutions. One compelling measure of institutional effectiveness is whether countries actually comply with their commitments. But compliance rates alone can be misleading. The possibility of selection bias make it difficult to draw conclusions about effectiveness from apparent compliance alone. Countries might comply because the institution is influential or because they simply selected international commitments they would have followed anyway in the absence of the institution (Downs and Rocke, 1995; Downs, Rocke, and Barsoom, 1996; Raustiala, 2000; Von Stein, 2005). It is often unclear what actions the country would have taken absent the international agreement. So it remains difficult to parse whether first-order compliance is due to the influence of the institution or merely the result of screening.

Second-order compliance provides a window into institutional effectiveness. Here, it is useful to evaluate changes in a country's behavior once it has violated its primary obligations—whether the country corrects an initial violation in response to some action by an international institution. International courts supply such a venue. Countries are routinely sued for alleged violations of their commitments and frequently found guilty. In these instances, a government's response to an adverse ruling can be

¹These arguments come in many flavors but are united by their use of game theoretic concepts to explain international cooperation, e.g. Stein (1990); Keohane (1982); Yarbrough and Yarbrough (1992); Goldsmith and Posner (2005).

quite revealing. By virtue of the initial breach and the guilty verdict, it is clear that a defendant government does not want to comply. But when, despite these preferences, it does comply, one can infer the influence of the international court. This situation gives the researcher leverage on the overarching question of effectiveness. By focusing on second-order compliance, this dissertation shows that international courts do prompt meaningful changes in state behavior. Defendant governments that lose international lawsuits do tend to comply with the court rulings in meaningful ways, adjusting their behavior to conform to the court's requirements. Evaluated in this way, the evidence shows that international courts have been successful.

Yet problems persist: the impact of international courts is conditional on domestic politics. When defendant governments face significant domestic political divisions, the initial breach of international law that provoked the lawsuit can remain locked into place, resulting in noncompliance. Domestic institutions and partisan politics pose substantial barriers that temper international courts' impact, and they do so in specific and predictable ways. Accordingly, the preceding chapters have examined international dispute settlement at two prominent international courts: the World Trade Organization's Dispute Settlement Mechanism and the European Court of Justice. The results show that adverse rulings from both international courts can prompt countries to follow through on their commitments to trade liberalization and economic integration, correcting the violations that provoked the disputes. Compliance varies with domestic politics within defendant governments. Defendants with many institutional constraints and partisan divisions are less likely to comply, and more likely to engage in extended legal battles with prolonged periods of defiance. I show that the deleterious impact of domestic constraints on compliance applies in both institutions, albeit more prominently under the WTO than the ECJ.

A central implication of this dissertation is that while international courts do shape government behavior in notable ways, they have not wrested power away from states. For both the WTO Dispute Settlement Mechanism and the European

Court of Justice, policymakers have expressed fear of judicial overreach, a concern shared by some scholars. The ECJ is often described as an agenda-driven institution that promotes EU integration, at times constraining the types of policy responses available to national governments and compromising their sovereignty (Alter, 1998; Burley and Mattli, 1993). Some studies have highlighted judicial activism at the ECJ but argue that it is tempered by the threat of coordinated noncompliance by member states (Garrett, Kelemen, and Schulz, 1998; Carrubba, Gabel, and Hankla, 2008). Similar questions have been raised about the WTO, although many scholars conclude that the Dispute Settlement Mechanism has steered clear of judicial activism and abides by principles of self-restraint (Steinberg, 2004; Jackson, 2006). The persistent influence of domestic politics, demonstrated above, implies that fears of judicial overreach are likely misplaced. Domestic institutions and preferences remain crucial and governments must make the final move to comply with or defy the rulings of international courts.

The persistent influence of domestic politics on international cooperation speaks to a broader dialogue about the legalization of international institutions (Goldstein and Martin, 2000; Koremenos, Lipson, and Snidal, 2001; Abbott and Snidal, 2000; Keohane, Moravcsik, and Slaughter, 2000). A key element of legalization is delegation, when states grant third parties the authority to implement, interpret and apply the rules, to resolve disputes and sometimes to make further rules (Abbott et al., 2000). By delegating authority to international courts, governments aim to insulate the dispute settlement process abroad from political controversy and strife at home (Tallberg, 2002; Reinhardt, 2002; Hawkins et al., 2006). Delegation provides some insurance that states will follow the legal principles embodied in international agreements rather than reacting to idiosyncratic domestic events. Yet in these two highly legalized international courts, domestic politics continues to play an important role, creating contingent successes and failures for dispute resolution. The findings above show that even when states delegate authority to international courts with

the intention to promote cooperation, those efforts may be thwarted at the domestic implementation stage.

5.2 Summary

This dissertation makes several contributions to the study of international organizations. The first empirical chapter argued that domestic veto players hinder compliance with adverse rulings from the WTO Dispute Settlement Mechanism. The analysis relied on a highly precise measure of *de facto* compliance based on bilateral product-level trade flows between the plaintiff and defendant countries in each WTO dispute. This chapter introduced a novel approach to measuring compliance using the synthetic control method.

The results showed that on average, countries with many domestic veto points were less likely to comply. By establishing this pattern as robust and probing the impact of veto players in several examples, this chapter helps reconcile disparate findings from previous studies on the impact of WTO rulings on trade. In short, WTO rulings do not have uniform effects on trade because the political conditions within defendant governments vary significantly. The finding that veto players hinder compliances has important bearings for understanding the relationship between domestic and international institutions, which I address below.

In the second empirical chapter, I tested the argument in the European Union, showing how judicial enforcement of economic integration is sensitive to domestic politics. Here, the main intervention of interest was an adverse ruling from the European Court of Justice in an infringement dispute over trade-related policies. European Union members that lost more lawsuits subsequently increased their imports from other members but this tendency depended on domestic politics. Countries with many domestic veto points were resistant to adverse rulings and their reliance on EU trade budged little, indicating noncompliance. These findings relied on a

hierarchical model that captured country-level and EU-wide variables important for understanding the impact of ECJ rulings on commerce. Even for the ECJ, a relatively rigid institution with well-developed enforcement mechanisms, domestic politics can hinder compliance.

The third empirical chapter evaluated the effectiveness of international dispute settlement along a different dimension: the time to resolve a dispute. In both the WTO and the ECJ, prompt dispute settlement is viewed as critical marker of success. But under both institutions, many disputes drag on for years unresolved. Because prolonged lawsuits can buy defendants time to “cheat” at the expense of plaintiffs and other members of the international organization, they can have deleterious effects on cooperation that are similar to noncompliance. This chapter demonstrated that lawsuits against defendants with many domestic constraints tended to last longer on average, before the countries acquiesced. Under both the WTO and ECJ, defendants with many veto points tended to delay and resist the court’s adverse ruling.

Furthermore, this chapter compared the two institutions, highlighting key ways in which the WTO has a more flexible design than the ECJ. I argued that more flexible designs can magnify the impact of domestic politics and showed that the ill-effects of veto players on dispute duration have been much stronger in the WTO than in the ECJ. In short, the design of the international court appears to mediate the impact of domestic veto players on dispute settlement.

5.3 Empirical Contributions and Theoretical Significance

This dissertation makes several empirical contributions. First, my argument speaks to the broader question of when international institutions are effective. Across the different paradigms in international relations there is broad agreement that the analytical obstacles to demonstrating institutional effects are substantial. Pervasive

selection bias, and the fact that experimental studies are nearly impossible, make causal inference difficult.

To gain leverage on this analytical puzzle, I focus on second-order compliance in the well-defined context of international court rulings. In both the European Union and the World Trade Organization, countries violate their primary commitments to economic integration and trade liberalization. These initial violations help to reveal state preferences. In ECJ disputes with an adverse ruling, the state's resistance to European economic integration tends to be quite strong, since otherwise it would have conformed to its legal obligations in earlier stages (Panke, 2007; Börzel, Hoffmann, and Panke, 2012). Similarly in the WTO, disputes that go through litigation and receive an adverse ruling are "tough cases" with obstinate defendants, else those governments would have corrected their violations sooner, saving the cost of litigation. So these lawsuits are clear instances in which the defendant government did not want to follow its primary obligations. When the defendant corrects the violation despite these preferences, one can infer the international institution had an impact. Second-order compliance supplies evidence that the international court affected state behavior by shifting incentives and making the violation untenable. By observing changes in countries' behavior in response to adverse rulings, researchers can draw more reliable conclusions about the effectiveness of international courts.

States surely draw inferences about institutional effectiveness by observing second-order compliance as well. The argument presented above suggests that international courts also play a demonstrative role in reinforcing international cooperation. When governments see defendants respond to adverse rulings by making policy changes with real economic effects, they may be further convinced that legalized dispute settlement works. By extension, they may be more willing to seek recourse within the international institution according to formal legal channels. The broader theoretical point is that second-order compliance, by conveying information about the effec-

tiveness of international dispute settlement, may enhance international cooperation among governments other than the disputants.

This dissertation has also highlighted the link between domestic political constraints countries face and their international behavior with respect to economic cooperation. Domestic institutional divisions and partisan opposition—veto points—create constraints that impede policy change. One compelling aspect of the veto points metric is that it subsumes multiple factors that other scholars have emphasized: proportional representation versus majoritarian systems, divided versus unified government, polarized versus homogeneous parties, etc. (Rogowski and Kayser, 2002; Helpman and Grossman, 2005; Kono, 2009; Evans, 2009; Rickard, 2010, 2012). These factors individually introduce minor obstacles to policy reform which, in aggregate, generate substantial hurdles with measurable effects. I highlight how these factors jointly shape states' international actions.

Multiple veto points in domestic government, it is widely argued, promote international cooperation. They do so by enhancing the credibility of countries' international commitments. Veto players narrow the set of potential agreements that a country can join because if any one veto player opposes the agreement, she can obstruct its adoption, for instance by blocking ratification. Therefore any international agreement that passes through the gauntlet of domestic politics represents a highly credible promise about future behavior. Moreover, once the international agreement is in place, domestic checks and balances or partisan divisions can make it difficult for the government to revise or revoke the agreement. For these reasons, scholars have concluded that democracies are much less likely to violate international agreements, once the agreements pass domestic scrutiny. Credibility is crucial in that it promotes the stable and efficient operation of the institution.

Unfortunately, the prevalence of lawsuits and guilty verdicts makes clear that governments with many domestic veto players can and do violate their international commitments. When a government has many domestic constraints and it violates its

primary international obligations, those violations also become difficult to reverse. Thus domestic veto players do not necessarily conduce to international cooperation; they promote policy continuity and stability, whether cooperative or conflictual. The key implication is that democracy, with its multiple veto points, isn't uniformly good for international cooperation. Sometimes core features of democratic politics do promote cooperation but other times they tilt the balance the opposite way, making it difficult to reverse breaches of international law and thereby obstructing international cooperation.

Another empirical contribution of this dissertation is its measurement of *de facto* compliance in WTO disputes. As discussed in Chapter 2, I use a synthetic control method to evaluate the impact of an adverse WTO ruling on product-level trade flows between disputing countries by creating an estimated counterfactual of what trade would have been in the absence of the ruling. By comparing actual trade to expected trade of disputed products in the wake of a WTO ruling, I infer whether the defendant government complied. This measurement strategy emphasizes the impact of a legal intervention on economic outcomes and reflects the WTO's broader trade-promoting agenda. It may be applicable to other studies of second-order compliance. Using this approach, I evaluated compliance in all WTO disputes where the defendant lost the ruling and was required to change its trade policy. In 35% of cases, I found strong evidence of compliance. But in nearly 45% of cases, trade in the disputed products continued to decline substantially, indicating noncompliance. These compliance measurements form the core of my WTO dispute dataset.

In addition to measuring compliance, the WTO dataset provides information about the legal process and litigants.² The vast majority (nearly 90%) of disputes that went through the litigation process resulted in an adverse ruling—where the de-

²Building on Horn and Mavroidis (2008), it includes the extent of the adverse legal ruling, the dates of each legal stage and corresponding durations, the number of third party countries that participated in the dispute and characteristics of the plaintiff and defendant countries.

fendant was found guilty of violating at least some of its international commitments. The average dispute with an adverse ruling lasted three years but some particularly intractable conflicts dragged on for nearly fifteen years. Over forty different countries have been sued under the WTO. They come from all regions of the world and display a range of domestic political conditions. Most lawsuits implicate wealthy industrialized countries but developing countries are increasingly willing and able to engage in WTO dispute settlement.

This dissertation also assembled data on infringement disputes at the European Court of Justice, building on Börzel and Knoll (2012). It includes information on infringement disputes on trade-related issues and the frequency of adverse rulings against EU member countries, and information on intra-European trade flows over time. EU members rely heavily on intra-EU commerce and almost all members draw well-over half of their imports from other members. By this measure, the EU is indeed tightly integrated. The data demonstrate wide variation in ECJ lawsuits. The countries of southern Europe are most frequently sued by the European Commission for infringing EU law on trade-related issues. The discrepancies across countries are substantial. Italy, for instance, was sued and found guilty of trade-related infringements sixty-two times during the sample period while Luxembourg was found guilty only twice. In sum, these data suggest that EU economic integration is deep but uneven and that many member governments remain reluctant participants, especially as the demands of EU integration have grown.

5.4 Questions and Implications

The findings in the preceding have implications for the optimal design of international institutions. There is now a rich literature that examines how an institution's design changes incentives and hence behavior of its members. One key design element is flexibility. When an international agreement is flexible, it is relatively permissive

of occasional defections and imposes only small penalties for violations. It makes partial compliance or settlement a viable option for member states under certain circumstances. More rigid agreements, by contrast, are less permissive of violations and impose larger penalties on members who breach the treaty terms.

Government leaders frequently experience domestic pressure to violate their international agreements. For example, they may be urged to violate a trade agreement by imposing protectionist policies that help politically influential constituencies. When domestic political pressure to violate is low, leaders may find it easy to comply with their international commitments. In these cases, a rigid treaty design makes full compliance preferable to settlement. But when a leader faces relatively high domestic pressure, rigid agreements—which make it difficult to temporarily violate and then settle—may force her to abandon the agreement. Therefore a rigid design tends to increase the probability of full compliance but decrease the stability of the international agreement. Previous literature has argued that flexibility can alleviate this risk and make an international agreement more stable (Rosendorff and Milner, 2001; Rosendorff, 2005; Johns, 2014; Johns and Peritz, 2015). Flexible designs, by allowing partial compliance in certain circumstances, reduce the likelihood that a government under domestic pressure defects from the agreement.

While most scholars have focused on flexibility and stability in a country's original decision whether to comply with its international commitments (first-order compliance), the analysis in this dissertation suggests a similar logic may hold for international dispute settlement (second-order compliance). Defendant governments in international disputes are sensitive to domestic politics. When a defendant government has many domestic political constraints, it will tend to resist adjusting its policies and domestic groups that oppose international cooperation can create pressure on government leaders. In the preceding chapters, I argued that a government faced with an adverse ruling from an international court is required to comply. But its actual compliance depends on the domestic constraints and political pressure it

faces. When political pressure is high, governments with many domestic constraints will be especially likely to resist the court.

Flexibility is likely to exacerbate this effect. A flexible court makes it easier for governments to resist an adverse ruling by delaying or reaching a negotiated settlement and reduces the risk that they will abandon the institution altogether. So when domestic political pressure is high, governments with many constraints will be most likely to use a court's flexibility by delaying, settling, or partially complying. A rigid court makes few such allowances. A highly constrained government that is under political pressure may find it costly to follow through on an adverse ruling. When the court is rigid, the government may find the costs to be intolerable and abandon the institution altogether. This suggests that an optimally-designed court should account for the extent of domestic political pressure and constraints that governments face. When member governments are highly constrained, a flexible international court may prompt countries to delay compliance with adverse rulings, or adopt only partially-compliant measures, but ensure greater stability over time.³

The preceding analysis also raises questions about whether international courts rule strategically and how domestic politics might matter for those decisions. The possibility of strategic rulings at the ECJ has been the topic of active scholarly debate (Carrubba, Gabel, and Hankla, 2008; Stone Sweet and Brunell, 2012; Carrubba, Gabel, and Hankla, 2012). Just as international courts may refrain from judgments that risk provoking coordinated backlash by members, judges may be hesitant to rule against defendant governments with many domestic constraints. Obstinate defendants that routinely defy court rulings can undermine the authority of the court because it occupies a somewhat precarious position of authority. An international court's legitimacy derives from the fact that member states delegate power and grant

³Political pressure is a function of idiosyncratic events and is difficult to anticipate while a government's domestic constraints are easier to predict. So it is more plausible that a court could be designed around countries' expectations about their domestic constraints.

them the authority to issue binding rulings (Abbott and Snidal, 2000; Abbott et al., 2000; Keohane, Moravcsik, and Slaughter, 2000; Hawkins et al., 2006; Buchanan and Keohane, 2006). This possibility has not received much attention in the current WTO literature. For both the ECJ and the WTO, it remains an open question whether judges issue rulings strategically and take into account the domestic political constraints that defendant governments face.⁴

The fact that international litigation does not consistently result in compliance opens additional questions about the purpose and function of international courts. *Why do plaintiffs sue defendant governments with many domestic veto players?* While some such lawsuits lead to compliance, many others are “bad bets” and, insofar as compliance is the goal, exercises in futility. This suggests that international courts play a role that is a great deal broader than providing a venue for enforcement. If international courts do not consistently induce compliance, why do countries use them?

Some scholars, adopting a managerialist perspective, have argued that international courts help inform countries about policy choices. According to this view, countries may be uncertain of whether a policy constitutes a violation and international courts can provide useful information (Fang et al., 2010; Hoekman and Mavroidis, 2000). With their rulings, international courts establish violations as such and specify a course of action to correct the violation. By informing states, courts promote cooperation. This assumes countries can and want to follow the rules—they only need better instruction. Unfortunately, this line of argument does

⁴For example, international courts could be wary of issuing adverse rulings that mandate wide-reaching reforms and bring to bear many veto players in a country. In both the WTO and ECJ, there are examples of such rulings. WTO disputes that implicate broad-reaching national legislation include lawsuits over the US Copyright Act or the US Anti-Dumping Act. In these instances, the United States was required to adopt new legislation to bring its policies into compliance, a process that activates many veto players. Similarly, several ECJ rulings have required policy reforms that rely on the consent of multiple veto players (e.g. the dispute over UK legislation on the origins of goods, *Commission v. United Kingdom*, C207/83.) Further analysis is needed to determine whether these types of rulings are exceptional or commonplace and implications for strategic rulings.

not satisfactorily explain the appearance of lawsuits against highly constrained defendants which, informed by the court's verdict, encounter too much domestic gridlock to respond.

Another explanation for the prevalence of lawsuits against obstinate defendants is that adjudication and principled assessment matters in itself. Insofar as governments value rule of law, international courts are a vehicle for promoting predictable and fair interactions between countries. Their judgments help coordinate countries and promote consensus about what types of behaviors are acceptable or unacceptable in the international economic arena. If governments observe these norms and internalize them over time, courts can play a crucial role in promoting international cooperation, even if the defendant is sometimes unable to comply.

Widening the analytical lens, a related explanation is that international lawsuits are part of repeated interactions between states over long periods of time. By suing a defendant government that faces many domestic constraints, a plaintiff may send a costly signal of its intolerance for rule violations. Even when it knows an adverse ruling may not be fully implemented, the plaintiff can draw attention to the violations and raise awareness in the international community of cheating behavior. Doing so may increase the likelihood that future violations by the same defendant will be collectively punished.

Exploring the relationship between domestic constraints and institutional design prompts additional questions about the broader environment in which international courts operate. When do countries choose more or less flexible international courts? Does this choice depend on their domestic constraints or their broader goals for international cooperation? International institutions are designed with a set of trade-offs and must function for a diverse set of countries with different domestic constraints and thus different propensities for compliance. As noted above, flexible designs can improve cooperation when member governments face significant domestic constraints while rigid designs may be more effective when governments are relatively

unconstrained. The WTO and ECJ represent different choices of institutional design that reflect these trade-offs. While these two courts play similar roles in dealing with trade disputes and enforcing trade liberalization, they operate in different political contexts with different ramifications.

The WTO is an institution with a tremendously diverse set of members with a wide range of domestic conditions. In this context, the choice of a flexibly designed dispute settlement mechanism made sense insofar as it provided for the varied challenges and conditions members could expect to face. Like escape clauses, a flexible DSM may have been a necessary precondition in initial negotiations in order to get consent from member states. And the WTO's compliance problems are a predictable consequence of the DSM's flexible design.

The European Court of Justice might too have been designed to be flexible since, in some domains of EU cooperation, we do see more flexible arrangements. For example, in monetary affairs, the EU has a system of *a la carte* integration where different countries choose whether or not to join the Eurozone at different times. But in trade-related aspects of EU integration, European member states adopted very deep forms of cooperation where temporary violations or partial compliance risked undermining the institution. Therefore, it is understandable that the ECJ was designed to be rigid and demand full compliance. A rigid design was possible because the ECJ enjoys stability provided by the rest of the EU institutions. For instance, by ignoring an ECJ ruling on trade, a member may incur penalties backed by the EU institutions that controls financial assistance and loans. Member governments cannot exit the EU without generating tremendous costs to their economy. A consequence is that compliance with ECJ rulings appears strong and there is little opportunity for domestic politics to sway outcomes.

In sum, institutional design choices are guided not only by countries' expectations about compliance and the compromises they must make to reach an agreement. They

also reflect the extent of integration—including the scope of cooperation across issue areas and policy domains—that countries are trying to achieve.⁵

This dissertation demonstrates that international courts can successfully restore or even expand economic cooperation between countries but their ability to do so hinges on domestic politics. The domestic political explanation I provide traces compliance with international legal rulings back to the institutional and partisan divisions within governments. On average, defendant governments that lose international disputes comply when they can overcome their domestic constraints. Despite the constant challenge of domestic politics, the evidence is compelling that international courts are indeed effective.

International economic cooperation remains a crucial contemporary challenge with real consequences not only for states but also for the workers, investors, and consumers who inhabit them. This dissertation has helped to illuminate the success of international dispute settlement and the cross-cutting role of domestic veto players in international cooperation. It has shown that even in the face of unexpected domestic challenges and diverse membership, international institutions can and do deliver on their promises.

⁵The extent of integration includes both breadth and depth of cooperation. This discussion has focused on breadth and suggests that there may be a relationship between the breadth of cooperation and the flexibility of the institution, including the design of its dispute settlement mechanism. In a recent study, Johns (2014) argues that there are trade-offs between the depth and rigidity in international trade agreements. The arguments explored above point to the need for further analysis, including whether broader agreements are more stable given rigid designs.

APPENDIX A

Technical Material for Analysis of WTO Disputes

A.1 Measuring Compliance with the Synthetic Control Method

Appendix A.1 describes the statistical approach used in Chapter 2, the synthetic control method. It draws directly on Abadie et al. (2012) and I use the ‘`synth`’ package in R.

The synthetic control method begins with a sample of units (i.e. countries, j) observed over multiple time periods (i.e. years, $t = 1, 2, \dots T$). The case of interest, the unit exposed to the intervention, is the “treated unit” ($j = 1$). Other units constitute the “donor pool”, the population of potential comparison units that may be used to approximate the counterfactual of the case of interest without the intervention ($j = 2, 3, \dots J$). The donor pool contains units with outcomes that are thought to be “driven by the same structural process as the unit representing the case of interest and that were not subject to structure shocks to the outcome variable during the sample period of the study” (Abadie et al. 2012, 4). The goal is to closely match the pre-intervention characteristics of the treated unit by using a combination of untreated units.

In this application, the “treated unit” is the respondent country that has received an adverse WTO ruling. The “donor pool” consists of countries other than the respondent. These countries are selected to be roughly similar to the respondent in terms of several economic characteristics. All countries are observed annually for

several years before the WTO dispute—the “pre-intervention period”—and several years after—the “post-intervention period.”

Since the donor pool contains several units, I average across them. The synthetic control is defined as the *weighted average* of units in the donor pool, represented by a column vector of weights,

$$\begin{aligned} W &= (w_2, \dots, w_{j+1})' \quad \text{wherein:} \\ 0 \leq w_j \leq 1 \quad &\text{for } j = 2, \dots, J \\ \text{and } w_2 + \dots + w_J &= 1. \end{aligned} \tag{A.1}$$

Weights are chosen to minimize the difference between the pre-intervention characteristics of the treated unit and the synthetic control. I refer to this vector W as the “unit-weights”.

In this instance, the pre-intervention characteristics are the variables that describe each country’s economy in the years leading up to the dispute. The variables include gross domestic product, unemployment rate, industry share of GDP, etc. as described above. I assign unit-weight to each donor pool country to ensure their unit-weighted average (the synthetic control) looks very much like the respondent country (the treated unit) in the years before the ruling.

To find the best vector of weights, the optimization process is as follows. Suppose there are k variables representing the different characteristics of interest. Let X_1 be a $(k \times 1)$ vector of values for the pre-intervention characteristics of the treated unit and X_0 be a $(k \times (J - 1))$ matrix with values for the same variables for units in the donor pool. The best vector of unit-weights, the synthetic control W^* , minimizes the size of the difference $X_1 - X_0W$. The optimization problem is to find the W^* that minimizes:

$$\|X_1 - X_0W\| V = \sqrt{(X_1 - X_0W)' V (X_1 - X_0W)} \quad (\text{A.2})$$

where V is a $(k \times k)$ symmetric and positive semidefinite matrix of weights for the variables in X_1 and X_0 . I refer to the matrix V as the “covariate-weight” matrix and describe it below. The vector of unit-weights W , reflects the relative importance assigned to each donor unit (i.e. country) when measuring the difference $X_1 - X_0W$. Countries with large predictive power on the outcome receive greater unit-weight.

For example, suppose the WTO rules against China in a dispute. I compare China to India, Japan, and the United States in the pre-intervention period. If China’s economic characteristics—GDP growth rate, unemployment rate, etc.—look more like India’s than Japan’s or the United States, then India would get a larger unit-weight. Which economic characteristics are most important in determining India’s similarity to China? To make this determination, the SCM also identifies optimal covariate-weights.

To compute the best covariate-weight matrix V^* , one must optimize over all reasonable covariate weighting schemes \mathcal{V} according to some optimization criteria. The ‘‘`synth`’’ package in R minimizes the mean squared prediction error (MSPE) of the outcome variable in the pre-intervention periods. Specifically, let Y_1^{pre} be the $(T_P \times 1)$ vector of values for the outcome variable for the treated unit in the pre-intervention periods. Note that $1 \leq T_P$ is the number of pre-intervention periods over which the MSPE is minimized. For example, if we have data for China during the ten years leading up to the adverse ruling, then $T_P = 10$ and the vector Y_1^{pre} is China’s trade in those ten years. Let Y_0^{pre} be the $(T_P \times (J - 1))$ analogous matrix for the $(J - 1)$ units in the donor pool. Using the example where the donor pool consists of India, Japan and the United States, the Y_0^{pre} matrix has ten rows and three columns. The values of the matrix are the countries’ yearly trade. These

data are used to compute an optimal covariate-weight matrix. So V^* is chosen to minimize:

$$\arg \min_{V \in \mathcal{V}} (Y_1^{pre} - Y_0^{pre} W^*(V))' (Y_1^{pre} - Y_0^{pre} W^*(V)). \quad (\text{A.3})$$

Note that \mathcal{V} is the set of all positive definite and diagonal matrices of covariate-weights. The unit-weights for the synthetic control are given by W^* as discussed above. Because the matrix V^* is positive diagonal, it ensures all covariates receive non-negative weights that directly predict the outcome variable and there are no covariate interactions.

In summary, the synthetic control method solves a nested optimization problem that minimizes equation A.3 for the $W^*(V)$ given by equation A.2. It selects an optimal vector of unit-weights W^* and an optimal matrix of covariate-weights V^* . The unit-weights W^* correspond to each country's contribution to the synthetic control unit and the covariate-weights V^* correspond to each covariate's contribution to the similarity determination.

Next, the synthetic control unit is used to estimate the treatment effect. The treatment effect is estimated by comparing post-intervention outcomes for the treated unit to those for the synthetic control unit (which is not exposed to the intervention). In the example above, we interpolate what China's trade *would have been* in the years following the WTO dispute had it not actually been "treated" with an adverse ruling. The interpolation is based on a combination of what India, Japan, and the US's actual trade in the post-dispute years. Since the vector of unit-weights W^* placed the most weight on India, the "synthetic-China" trade pattern looks a lot like India's actual trade pattern. I compare China's *actual* trade and the synthetic-China *interpolated* trade to estimate the treatment effect.

To calculate the treatment effect, I compare the post-intervention outcomes for the treated unit to the synthetic control unit for every year in the post-intervention

period. Let Y_{jt}^{post} be the outcome of unit j at time t . So Y_{1t}^{post} is the vector of post-intervention values of the outcome for the treated unit. In the example, this is China's trade in the post-dispute years. The matrix of post-intervention values of the outcome for the donor pool is denoted Y_{jt}^{post} where $j = 2, \dots, J$. The matrix Y_{jt}^{post} contains India, Japan and US trade in the post-dispute years.

The synthetic control estimator of the effect of the treatment is given by:

$$\hat{\alpha}_{1t} = Y_{1t}^{post} - \sum_{j=2}^J w_j^* Y_{jt}^{post} \quad (\text{A.4})$$

(Abadie and Gardeazabal 2003; Abadie, Diamond and Hainmueller 2010). This is the difference between the vector of post-intervention values for the treated unit minus the unit-weighted average of post-intervention values for the donor pool. In summary, this technique establishes a synthetic control unit that has similar behavior to the case of interest during the time period prior to the intervention and interprets discrepancy in the outcome variable following the intervention as produced by the intervention itself. Compliance is then measured with a summary statistic and associated uncertainty. Denote the time of treatment to be $t = \tilde{t}$. Recall the synthetic control estimator of the effect of the treatment is given by equation A.4 for the post-treatment period, $t > \tilde{t}$. Similarly, I define a goodness-of-fit estimator for the pre-treatment period by:

$$\hat{\alpha}_{1t}^{pre} = Y_{1t}^{pre} - \sum_{j=2}^J w_j^* Y_{jt}^{pre} \quad \text{for } t \leq \tilde{t}. \quad (\text{A.5})$$

The compliance score S is calculated as the difference in means. This is the average effect of the treatment in the post-intervention period minus the average goodness-of-fit in the pre-intervention period:

$$S = \frac{1}{(T - \tilde{t})} \sum_{\tilde{t}+1}^T \hat{\alpha}_{1t} - \frac{1}{\tilde{t}} \sum_1^{\tilde{t}} \hat{\alpha}_{1t}^{pre}. \quad (\text{A.6})$$

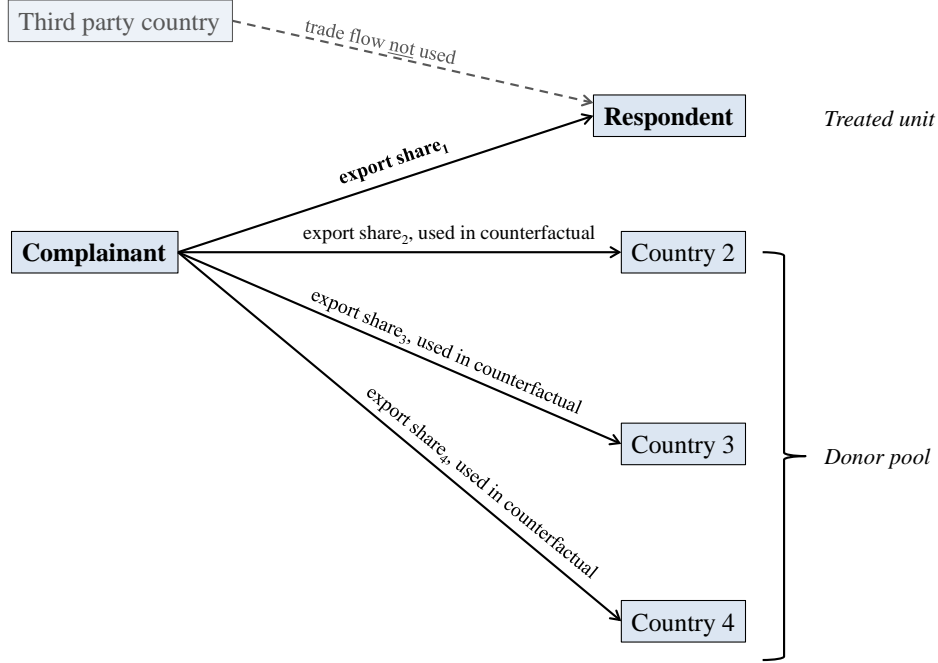
The standard deviation is computed for the $\hat{\alpha}_{1t}^{pre}$ values. This standard deviation denotes, in a sense, how well the synthetic control method performed in the optimization calculations. If the synthetic control unit was very similar to the treated unit in all years of the pre-intervention period, then the two should vary over time together and the standard deviation will be small. The compliance score should be quite trustworthy. If the synthetic control unit, on the other hand, is very dissimilar to the treated unit in some of the pre-intervention years, then the standard deviation will be large. When the standard deviation is large, the compliance score is less reliable. Thus together the compliance score and standard deviation capture the extent of compliance and the reliability of the metric.

A.2 Trade Flows Used in Synthetic Control Method

When third party countries enter WTO disputes, they most often side with the complainant country. The trade flows used to create counterfactual are not the same as the trade flows at stake when third party countries enter WTO disputes. Figure A2.1 shows a schematic of the trade flows used and not used to create the counterfactual.

Rarely, third party countries side with the respondent. In these cases, including the third parties in the counterfactual could bias my measurement of compliance. To guard against this problem, I exclude third party countries whenever sufficient data are available. Regardless, this form of measurement bias should not be systematically correlated with the respondent country's domestic veto players.

Figure A.1: Schematic of Trade Flows



Note: Exports from the Complainant to other countries are used to create the counterfactual. Third party countries export to the Respondent but these trade flows are not used.

A.3 Robustness Checks for Selection Effects

To check for possible selection bias, I use a Heckman selection model. The first stage (selection equation) indicates the probability that the governments fail to resolve their dispute during consultations and go on to receive an adverse ruling. Because almost all cases that go to litigation receive some form of adverse ruling, the selection stage closely approximates selection into litigation. The second stage (outcome equation) indicates the probability that respondents comply, conditional on going to litigation and receiving an adverse ruling.

The main difficulty with selection models is achieving identification which requires a variable that predicts selection but is not correlated with the outcome (Heckman, 1979, 1990; Dubin and Rivers, 1989). I aim to identify the model with several variables used in related studies (Przeworski and Vreeland, 2000; Von Stein, 2005; Busch and Pelc, 2010; Johns and Pelc, 2014). I use indicators for agricultural disputes, health and safety disputes, and disputes over general trade laws (e.g. US Anti-Dumping Act of 1916, US tax treatment for “foreign sales corporations”). Disputes over general trade laws are frequently litigated but are not consistently correlated with compliance. I also include the exchange rate between the complainant and respondent and their similarity in political preferences, measured as the distance between their United Nations General Assembly (UNGA) vote ideal point estimates (Bailey, Strezhnev, and Voeten, 2013). The greater the distance between the disputants’ ideal points, the more likely they are to litigate and the less likely to resolve their disagreement through negotiated settlement. The similarity between the disputants’ ideal points is not correlated with compliance because once the WTO issues a legal verdict, the burden for action is largely the respondent’s alone. Together, these variables serve as a plausible, if imperfect, source of identification. I also control for the market size of the complainant and respondent, an indicator for anti-dumping disputes, the number of third parties, and cases where the EU is the respondent.

Table A.1 shows the results. In the selection stage, both respondents and complainants with larger economies are less likely to settle and more likely to litigate. This suggests larger countries are better able to bear the costs of litigation and “hold out” for a more favorable outcome than they could obtain during negotiations. The tendency is pronounced for the respondents: when economically powerful, they tend not to back down before receiving a legal verdict. The correlation is not driven by the EU because, while Europe litigates more frequently than other respondents, the tendency is not statistically significant. The number of third parties also reduces the probability of settlement.

Veto points in the respondent government are not consistently associated with early settlement versus litigation. Compared to litigation, the consultation process shields negotiators from demands at home and allows governments to settle early—or not—without being beholden to domestic political groups. Some research suggests that WTO consultations allow governments privacy and discretion in resolving disputes, a feature that is especially important for dispute resolution between democracies. The topic of the dispute—agriculture or health and safety—is not a significant predictor of litigation. Nor is the exchange rate or the distance between their UNGA voting ideal points. Disputes over general trade laws, which impact most or all products, are almost always litigated.

In the outcome stage, respondents with more veto points are less likely to comply. The extent of the adverse ruling tends to predict compliance, with some model specifications indicating a statistically significant positive correlation. International pressure, represented by the number of third parties, is associated with more compliance. Bearing in mind the limitations, the Heckman selection models lend further support to the hypothesis.¹

¹ There are other selection forces. The observed WTO disputes are a subset of potential cases. Potential complainants should be able to (albeit imperfectly) anticipate the likelihood that respondent governments will comply with an adverse ruling. With some probability, potential complainants select out of the entire dispute process when their trade conflicts are against states with many domestic veto players. Forgoing the dispute may be preferable to initiating a costly and futile legal battle.

This form of selection does not undermine the broader argument. Governments that violate their primary WTO obligations are unlikely to come into compliance of their own accord. By “cheating” with an illegal trade barrier, that government gains competitive advantages, and has every incentive to keep the policy until challenged. In the absence of a lawsuit, potential respondents with many veto players should prolong their violations indefinitely. The negative link between domestic veto players and international cooperation may even be stronger than my results suggest.

Table A.1: Two-Step Heckman Selection Models for WTO Disputes, 1995-2011

Selection Equation: Is there an adverse ruling?					
	(1)	(2)	(3)	(4)	(5)
Veto Points	−0.071 (0.598)	0.636 (0.648)	0.636 (0.648)	0.677 (0.641)	−0.071 (0.598)
Respondent GDP	0.390*** (0.079)	0.270*** (0.096)	0.270*** (0.096)		0.390*** (0.079)
Complainant GDP	0.177** (0.075)	0.127 (0.096)	0.127 (0.096)		0.177** (0.075)
EU Respondent	0.185 (0.196)	0.108 (0.204)	0.108 (0.204)	0.004 (0.196)	0.185 (0.196)
Antidumping Dispute		0.339* (0.194)	0.339* (0.194)	0.377** (0.190)	
Agriculture Dispute		0.003 (0.222)	0.003 (0.222)	−0.013 (0.219)	
Health & Safety Dispute		−0.086 (0.276)	−0.086 (0.276)	−0.120 (0.271)	
All Products		1.084*** (0.387)	1.084*** (0.387)	1.201*** (0.377)	
Exchange Rate		0.00001 (0.00004)	0.00001 (0.00004)	0.00001 (0.00004)	
Ideal Point Distance		0.063 (0.095)	0.063 (0.095)	0.200*** (0.073)	
Third Parties		0.168*** (0.052)	0.168*** (0.052)	0.187*** (0.052)	
Constant	−0.313 (0.254)	−0.902*** (0.344)	−0.902*** (0.344)	−1.140*** (0.328)	−0.313 (0.254)
Outcome Equation: Did the respondent comply?					
	(1)	(2)	(3)	(4)	(5)
Veto Points	−0.842** (0.370)	−0.862** (0.369)	−0.901** (0.384)	−0.861** (0.378)	−1.325** (0.670)
Third Parties	0.047** (0.019)	0.035 (0.022)	0.034 (0.026)	0.035 (0.023)	0.093*** (0.035)
% Adverse	0.276** (0.123)	0.302** (0.126)	0.311** (0.128)	0.303** (0.126)	0.380* (0.225)
Respondent GDP			−0.002 (0.066)	0.021 (0.046)	
Complainant GDP			−0.020 (0.051)	−0.011 (0.046)	
Constant	0.685*** (0.210)	0.718*** (0.199)	0.744*** (0.283)	0.704*** (0.235)	1.146*** (0.381)
Compliance Measure	(0/1)	(0/1)	(0/1)	(0/1)	(0/1/2)
Total Observations	330	324	324	324	330
Outcome Observations	125	125	125	125	125
R ²	0.122	0.129	0.131	0.132	0.105

Note: Compliance was coded using the synthetic control method with annual bilateral trade data for disputed products. Estimates are calculated with a generalized tobit model and "sampleSelection" package. The outcome stage estimates the effect of covariates on compliance, conditional on the dispute receiving a ruling. *p<0.1; **p<0.05; ***p<0.01.

APPENDIX B

Technical Material for Analysis of ECJ Disputes

B.1 Code for Bayesian Analysis

This Appendix provides the Bugs code used in Chapter 4.

```
# Simple Model
model{
  for(i in 1:n){
    y[i] ~ dnorm(y.hat[i], tau.y)
    y.hat[i] <- alpha + inprod(beta[],X[i,])
  }
  tau.y <- pow(sigma.y, -2)
  sigma.y ~ dunif(0.1,100)
  alpha ~ dnorm(0,0.01)
  beta ~ dmnorm(mu.b, tau.b)
}

# Random Intercepts Model
model{
  for(i in 1:n){
    y[i] ~ dnorm(y.hat[i], tau.y)
    y.hat[i] <- alpha[cty[i]] + inprod(beta[],X[i,])
  }
  beta ~ dmnorm(mu.b, tau.b)
  tau.y <- pow(sigma.y, -2)
  sigma.y ~ dunif(0,100)
  for(j in 1:J){
    alpha[j] ~ dnorm(mu.a, tau.a)
  }
}
```

```

mu.a ~ dnorm(0,0.0001)
tau.a <- pow(sigma.a, -2)
sigma.a ~ dunif(0,100)
}

# Random Slopes Model
model {
  for (i in 1:n){
    y[i] ~ dnorm(y.hat[i], tau.y)
    y.hat[i] <- a + g[cty[i]]*Z[i] + inprod(b[],X[i,])
  }
  b ~ dmnorm(mu.b,tau.b)
  a ~ dnorm(0, 0.0001)
  tau.y <- pow(sigma.y, -2)
  sigma.y ~ dunif(0,100)
  for (j in 1:J){
    g[j] ~ dnorm(mu.g, tau.g)
  }
  mu.g ~ dnorm(0,0.0001)
  tau.g <- pow(sigma.g, -2)
  sigma.g ~ dunif(0, 100)
}

# Random Intercepts and Slopes Model
model {
  for (i in 1:n){
    y[i] ~ dnorm(y.hat[i], tau.y)
    y.hat[i] <- a[cty[i]] + g[cty[i]]*Z[i] + inprod(b[],X[i,])
  }
  b ~ dmnorm(mu.b,tau.b)
  tau.y <- pow(sigma.y, -2)
  sigma.y ~ dunif(0,100)
  for (j in 1:J){
    a[j] ~ dnorm(mu.a, tau.a)
    g[j] ~ dnorm(mu.g, tau.g)
  }
  mu.a ~ dnorm(0,0.0001)
  tau.a <- pow(sigma.a, -2)
  sigma.a ~ dunif(0, 100)
}

```

```
mu.g ~ dnorm(0,0.0001)
tau.g <- pow(sigma.a, -2)
sigma.g ~ dunif(0, 100)
}
```

BIBLIOGRAPHY

- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105(490).
- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2011. "Synth: An R Package for Synthetic Control Methods in Comparative Case Studies." *American Statistical Association* .
- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2014. "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science* Forthcoming.
- Abadie, Alberto, and Javier Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review* 93(1): 113–132.
- Abbott, Kenneth, and Duncan Snidal. 1998. "Why States Act Through Formal International Organizations." *Journal of Conflict Resolution* 42(1): 3–32.
- Abbott, Kenneth W., and Duncan Snidal. 2000. "Hard and Soft Law in International Governance." *International Organization* 54(3): 421–456.
- Abbott, Kenneth W., Robert O. Keohane, Andrew Moravcsik, Anne-Marie Slaughter, and Duncan Snidal. 2000. "The Concept of Legalization." *International Organization* 54(3): 401–419.
- Allee, Todd L., and Paul K. Huth. 2006. "Legitimizing Dispute Settlement: International Legal Rulings as Domestic Political Cover." *American Political Science Review* 100(2): 219–234.
- Alter, Karen J. 1998. "Who are the "Masters of the Treaty"? European governments and the European Court of Justice." *International Organization* 52(1): 121–147.

- Alter, Karen J. 2001. *Establishing the Supremacy of European Law: The Making of an International Rule of Law in Europe*. Oxford: Oxford University Press.
- Alter, Karen J., and Laurence R. Helfer. 2010. "Nature or nurture? Judicial Law-making in the European Court of Justice and the Andean Tribunal of Justice." *International Organization* 64(4): 563–592.
- Alter, Karen J., Laurence R. Helfer, and Osvaldo Saldias. 2012. "Transplanting the European Court of Justice: the experience of the Andean Tribunal of Justice." *American Journal of Comparative Law* 60(3): 629–664.
- Bailey, Michael, Anton Strezhnev, and Erik Voeten. 2013. "Estimating Dynamic State Preferences from United Nations Voting Data." *Available from: <http://hdl.handle.net/1902.1/12379>*.
- Bechtel, Michael M., and Thomas Sattler. 2015. "What is Litigation in the World Trade Organization Worth?" *International Organization* 69(2): 375–403.
- Beck, Thorsten, George Clarke, Alberto Groff, Phillip Keefer, and Patrick Walsh. 2001. "New Tools in Comparative Political Economy: The Database of Political Institutions." *World Bank Economic Review* 15(1): 165–176.
- Bednar, Jenna. 2009. *The Robust Federation: Principles of Design*. Cambridge: Cambridge University Press.
- Börzel, Tanja A. 2001. "Non-Compliance in the European Union: Pathology or Statistical Artefact?" *Journal of European Public Policy* 8(5): 803–824.
- Börzel, Tanja A, and Moritz Knoll. 2012. "Quantifying Non-Compliance in the EU: A Database on EU Infringement Proceedings." *Available: <http://www.infringement-db.jmce.org/Index.html>* Freie Universität Berlin.

- Börzel, Tanja A., Tobias Hofmann, and Diana Panke. 2012. "Caving In or Sitting It Out? Longitudinal Patterns of Non-Compliance in the European Union." *Journal of European Public Policy* 19(4): 37–41.
- Börzel, Tanja A., Tobias Hofmann, Diana Panke, and Carina Sprungk. 2010. "Obstinate and Inefficient: Why Member States Do Not Comply With European Law." *Comparative Political Studies* 43(11): 1363–1390.
- Bown, Chad P. 2002. "The Economics of Trade Disputes, the GATT's Article XXIII, and the WTO's Dispute Settlement Understanding." *Economics & Politics* 14(3): 283–323.
- Bown, Chad P. 2004a. "On the Economic Success of GATT/WTO Dispute Settlement." *Review of Economics and Statistics* 86(3): 811–823.
- Bown, Chad P. 2004b. "Trade Disputes and the Implementation of Protection Under the GATT: An Empirical Assessment." *Journal of International Economics* 62(2): 263–294.
- Bown, Chad P. 2005. "Participation in WTO Dispute Settlement: Complainants, Interested Parties, and Free Riders." *The World Bank Economic Review* 19(2): 287–310.
- Bown, Chad P., and Kara M. Reynolds. 2014. "Trade Flows and Trade Disputes." *World Bank Policy Research Working Paper* (6979).
- Bown, Chad P., and Kara M. Reynolds. 2015. "Trade Agreements and Enforcement: Evidence from WTO Dispute Settlement." *World Bank Policy Research Working Paper* (7242).
- Brewster, Rachel. 2011. "The Remedy Gap: Institutional Design, Retaliation, and Trade Law Enforcement." *George Washington Law Review* 80: 102.

- Buchanan, Allen, and Robert O Keohane. 2006. "The legitimacy of global governance institutions." *Ethics & international affairs* 20(4): 405–437.
- Burley, Anne-Marie, and Walter Mattli. 1993. "Europe before the Court: a political theory of legal integration." *International organization* 47(01): 41–76.
- Busch, Marc L, and Eric Reinhardt. 2000. "Bargaining in the Shadow of the Law: Early Settlement in GATT/WTO Disputes." *Fordham International Law Journal* 24: 158.
- Busch, Marc L, and Krzysztof J Pelc. 2010. "The politics of judicial economy at the World Trade Organization." *International Organization* 64(2): 257–279.
- Carrubba, Clifford J. 2005. "Courts and Compliance in International Regulatory Regimes." *Journal of Politics* 67(3): 669–689.
- Carrubba, Clifford J, and Lacey Murrah. 2005. "Legal integration and use of the preliminary ruling process in the European Union." *International Organization* 59(02): 399–418.
- Carrubba, Clifford J., and Matthew Gabel. 2015. *International Courts and the Performance of International Agreements*. Cambridge University Press.
- Carrubba, Clifford J., Matthew Gabel, and Charles Hankla. 2008. "Judicial Behavior Under Political Constraints: Evidence from the European Court of Justice." *American Political Science Review* 102(11): 435–452.
- Carrubba, Clifford J, Matthew Gabel, and Charles Hankla. 2012. "Understanding the role of the European Court of Justice in European integration." *American Political Science Review* 106(01): 214–223.
- Chaudoin, Stephen, Jeffrey Kucik, and Krzysztof Pelc. 2013. "Do WTO Disputes Actually Increase Trade?". In *American Political Science Association 2013 Annual Meeting*.

- Chayes, Abram, and Antonia Handler Chayes. 1993. "On Compliance." *International Organization* 47(02): 175–205.
- Cox, David Roxbee, and E Joyce Snell. 1989. *Analysis of binary data*. Vol. 32 CRC Press.
- Dai, Xinyuan. 2005. "Why Comply? The Domestic Constituency Mechanism." *International Organization* 59(02): 363–398.
- Davis, Christina L. 2012. *Why Adjudicate? Enforcing Trade Rules in the WTO*. Princeton: Princeton University Press.
- Downs, George W, and David M Roche. 1995. *Optimal imperfection?: Domestic uncertainty and institutions in international relations*. Princeton University Press Princeton, NJ.
- Downs, George W., David M. Roche, and Peter N. Barsoom. 1996. "Is the Good News About Compliance Good News About Cooperation?" *International Organization* 50(3): 379–406.
- Dubin, Jeffrey A, and Douglas Rivers. 1989. "Selection bias in linear regression, logit and probit models." *Sociological Methods & Research* 18(2-3): 360–390.
- EuroStat. 2014. "International Trade Database ComExt." *Available from: <http://epp.eurostat.ec.europa.eu>* .
- Evans, Carolyn L. 2009. "A protectionist bias in majoritarian politics: An empirical investigation." *Economics & Politics* 21(2): 278–307.
- Fang, Songying et al. 2010. "The strategic use of international institutions in dispute settlement." *Quarterly Journal of Political Science* 5(2): 107–131.
- Finnemore, Martha. 1993. "International Organizations as Teachers of Norms: the United Nations Educational, Scientific, and Cultural Organization and Science Policy." *International Organization* 47(04): 565–597.

- Finnemore, Martha, and Kathryn Sikkink. 1998. "International Norm Dynamics and Political Change." *International Organization* 52(04): 887–917.
- Fisher, Roger. 1981. *Improving Compliance with International Law*. Charlottesville: University Press of Virginia.
- Gabel, Matthew J., Clifford J. Carrubba, Caitlin Ainsley, and Donald M. Beaudette. 2012. "Of Courts and Commerce." *The Journal of Politics* 74: 1125–1137.
- Garrett, Geoffrey, R Daniel Kelemen, and Heiner Schulz. 1998. "The European Court of Justice, national governments, and legal integration in the European Union." *International Organization* 52(1): 149–176.
- Gilligan, Michael J. 1997. *Empowering Exporters: Reciprocity, Delegation, and Collective Action in American Trade Policy*. Ann Arbor: University of Michigan Press.
- Gilligan, Michael J. 2004. "Is There a Broader-Deeper Trade-off in International Multilateral Agreements?" *International Organization* 58(3): 459–484.
- Gilligan, Michael, Leslie Johns, and B. Peter Rosendorff. 2010. "Strengthening International Courts and the Early Settlement of Disputes." *Journal of Conflict Resolution* 54(1): 5–38.
- Goldsmith, Jack L., and Eric A. Posner. 2005. *The Limits of International Law*. Vol. 199 Oxford: Oxford University Press.
- Goldstein, Judith, and Lisa L. Martin. 2000. "Legalization, trade liberalization, and domestic politics: a cautionary note." *International Organization* 54(3): 603–632.
- Goldstein, Judith, and Richard Steinberg. 2008. "Negotiate or Litigate? Effects of WTO Judicial Delegation on US Trade Politics." *Law and Contemporary Problems* pp. 257–282.

- Goldstein, Judith, Douglas Rivers, and Michael Tomz. 2007. "Institutions in International Relations: Understanding the Effects of the GATT and the WTO on World Trade." *International Organization* 61(1): 37–67.
- Gowa, Joanne, and Soo Yeon Kim. 2005. "An exclusive country club: the effects of the GATT on trade, 1950–94." *World Politics* 57(04): 453–478.
- Hafner-Burton, Emilie M., Jana von Stein, and Erik Gartzke. 2008. "International Organizations Count: What Statistics Tell Us About IOs." *Journal of Conflict Resolution* 52(2).
- Hawkins, Darren G., David A. Lake, Daniel L. Nielson, and Michael J. Tierney. 2006. "Delegation Under Anarchy: States, International Organizations, and Principal-Agent Theory." In *Delegation and Agency in International Organizations*, ed. Darren G. Hawkins, David A. Lake, Daniel L. Nielson, and Michael J. Tierney. Cambridge: Cambridge University Press pp. 3–38.
- Heckman, James. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* pp. 153–161.
- Heckman, James. 1990. "Varieties of selection bias." *The American Economic Review* 80(2): 313–318.
- Helpman, Elhanan, and Gene M Grossman. 2005. "A Protectionist Bias in Majoritarian Politics." *Quarterly journal of economics* (4): 1239–1282.
- Henisz, Witold J. 2000. "The Institutional Environment for Economic Growth." *Economics & Politics* 12(1): 1–31.
- Henisz, Witold J. 2002. "The Political Constraint Index (POLCON) Dataset." *Available: <http://www-management.wharton.upenn.edu/henisz/>*.

- Henisz, Witold J., and Edward D. Mansfield. 2006. "Votes and Vetoes: the Political Determinants of Commercial Openness." *International Studies Quarterly* 50(1): 189–212.
- Hiscox, Michael J. 2006. "Through a Glass and Darkly: Attitudes Toward International Trade and the Curious Effects of Issue Framing." *International Organization* 60(7): 755–780.
- Hoekman, Bernard M, and Petros Mavroidis. 2000. "WTO Dispute Settlement, Transparency and Surveillance." *The World Economy* 23(4): 527–542.
- Horn, Henrik, and Petros C. Mavroidis. 2008. "The WTO Dispute Settlement Data Set 1995-2006." Available from: <http://go.worldbank.org/X5EZPHXJY0> .
- Hudec, Robert E. 1999. "The New WTO Dispute Settlement Procedure: An Overview of the First Three Years." *Minnesota Journal of Global Trade* 8: 1.
- Imai, Kosuke, and Marc Ratkovic. 2014. "Covariate Balancing Propensity Score." *Journal of the Royal Statistical Society* 76(1): 243–263.
- Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. West Sussex, UK: John Wiley & Sons, Ltd.
- Jackson, John H. 2006. *Sovereignty, the WTO, and Changing Fundamentals of International Law*. Vol. 18 Cambridge: Cambridge University Press.
- Johns, Leslie. 2011. "Courts as Coordinators: Endogenous Enforcement and Jurisdiction in International Adjudication." *Journal of Conflict Resolution* 56(2): 257–289.
- Johns, Leslie. 2014. "Depth versus Rigidity in the Design of International Trade Agreements." *Journal of Theoretical Politics* 26(3): 468–495.
- Johns, Leslie. 2015. *Strengthening International Courts: The Hidden Costs of Legalization*. Ann Arbor, MI: University of Michigan Press.

- Johns, Leslie, and B. Peter Rosendorff. 2009. "Dispute Settlement, Compliance, and Domestic Politics." In *Frontiers of Economics and Globalization*, ed. James Hartigan. Emerald Group Publishing Limited pp. 139–163.
- Johns, Leslie, and Krzysztof J. Pelc. 2014. "Who Gets to Be in the Room? Manipulating Participation in WTO Disputes." *International Organization* 68(3): 663–699.
- Johns, Leslie, and Lauren Peritz. 2015. "The Design of Trade Agreements." In *The Oxford Handbook of the Political Economy of International Trade*, ed. Lisa Martin. Oxford University Press pp. 337–359.
- Keefer, Philip, and David Stasavage. 2003. "The limits of delegation: Veto players, central bank independence, and the credibility of monetary policy." *American Political Science Review* 97(03): 407–423.
- Keohane, Robert O. 1982. "The Demand for International Regimes." *International Organization* 36(02): 325–355.
- Keohane, Robert O, Andrew Moravcsik, and Anne-Marie Slaughter. 2000. "Legalized dispute resolution: Interstate and transnational." *International organization* 54(03): 457–488.
- Kono, Daniel Y. 2006. "Optimal Obfuscation: Democracy and Trade Policy Transparency." *American Political Science Review* 100(3): 369–384.
- Kono, Daniel Yuichi. 2009. "Market structure, electoral institutions, and trade policy." *International Studies Quarterly* 53(4): 885–906.
- Koremenos, Barbara, Charles Lipson, and Duncan Snidal. 2001. "The Rational Design of International Institutions." *International Organization* 55(4): 761–799.

- Kucik, Jeffrey, and Eric Reinhardt. 2008. "Does Flexibility Promote Cooperation? An Application to the Global Trade Regime." *International Organization* 62(3): 477–505.
- Lohmann, Susanne, and Sharyn O'Halloran. 1994. "Divided Government and US Trade Policy: Theory and Evidence." *International Organization* 48(4): 595–632.
- Mansfield, Edward D., and Helen V. Milner. 2012. *Votes, Vetoes, and the Political Economy of International Trade Agreements*. Princeton: Princeton University Press.
- Mansfield, Edward D., Helen V. Milner, and B. Peter Rosendorff. 2002. "Why Democracies Cooperate More: Electoral Control and International Trade Agreements." *International Organization* 56(3): 477–513.
- Mansfield, Edward D., Helen V. Milner, and Jon C. Pevehouse. 2007. "Vetoing Cooperation: The Impact of Veto Players on Preferential Trading Arrangements." *British Journal of Political Science* 37: 403–432.
- Marshall, Monty G., and Keith Jagers. 2012. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2012."
- Martin, Lisa. 2013. "Against Compliance." In *Interdisciplinary Perspectives on International Law and International Relations: The State of the Art*, ed. Jeffrey L. Dunoff, and Mark A. Pollack. Cambridge University Press pp. 591–610.
- Martin, Lisa L. 2000. *Democratic Commitments: Legislatures and International Cooperation*. Princeton: Princeton University Press.
- Mbaye, Heather AD. 2001. "Why National States Comply with Supranational Law Explaining Implementation Infringements in the European Union, 1972-1993." *European Union Politics* 2(3): 259–281.

- Mearsheimer, John J. 1994. "The False Promise of International Institutions." *International Security* pp. 5–49.
- Milner, Helen V., and B. Peter Rosendorff. 1996. "Trade Negotiations, Information and Domestic Politics: The Role of Domestic Groups." *Economics & Politics* 8(2): 145–189.
- Milner, Helen V., 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(1): 117–146.
- Panke, Diana. 2007. "The European Court of Justice as an agent of Europeanization? Restoring compliance with EU law." *Journal of European Public Policy* 14(6): 847–866.
- Pelc, Krzysztof. 2009. "Seeking Escape: The Use of Escape Clauses in International Trade Agreements." *International Studies Quarterly* 53(2): 349–368.
- Pelc, Krzysztof J. 2014. "The Politics of Precedent in International Law: A Social Network Application." *American Political Science Review* 108(03): 547–564.
- Phelan, William. 2015. *In Place of Inter-State Retaliation: The European Union's Rejection of WTO-style Trade Sanctions and Trade Remedies*. Cambridge: Cambridge University Press.
- Przeworski, Adam, and James Raymond Vreeland. 2000. "The effect of IMF programs on economic growth." *Journal of development Economics* 62(2): 385–421.
- Raustiala, Kal. 2000. "Compliance and Effectiveness in International Regulatory Cooperation." *Case Western Reserve Journal of International Law* 32: 387.
- Reinhardt, Eric. 2001. "Adjudication Without Enforcement in GATT Disputes." *Journal of Conflict Resolution* 45(2): 174–195.

- Reinhardt, Eric. 2002. "Tying Hands without a Rope: Rational Domestic Response to International Institutional Constraints." In *Locating the Proper Authorities: The Interaction of Domestic and International Institutions*, ed. Daniel A. Drezner. Ann Arbor: University of Michigan Press pp. 77–104.
- Rickard, Stephanie J. 2010. "Democratic Differences: Electoral Institutions and Compliance with GATT/WTO Agreements." *European Journal of International Relations* 16(4): 711–729.
- Rickard, Stephanie J. 2012. "A Non-Tariff Protectionist Bias in Majoritarian Politics: Government Subsidies and Electoral Institutions." *International Studies Quarterly* 56(4): 777–785.
- Rogowski, Ronald. 1989. *Commerce and Coalitions: How Trade Affects Domestic Political Arrangements*. Princeton, NJ: Princeton University Press.
- Rogowski, Ronald, and Mark Andreas Kayser. 2002. "Majoritarian electoral systems and consumer power: price-level evidence from the OECD countries." *American Journal of Political Science* pp. 526–539.
- Rose, Andrew. 2004. "Do We Really Know that the WTO Increases Trade?" *American Economic Review* 94(1): 98–114.
- Rosendorff, B. Peter. 2005. "Stability and Rigidity: Politics and Design of the WTO's Dispute Settlement Procedure." *American Political Science Review* 99(3): 389–400.
- Rosendorff, B. Peter, and Helen V. Milner. 2001. "The Optimal Design of International Trade Institutions: Uncertainty and Escape." *International Organization* 55(4): 829–857.
- Rubin, Donald B. 2005. "Causal Inference Using Potential Outcomes." *Journal of the American Statistical Association* 100(469).

- Simmons, Beth A. 1998. "Compliance with International Agreements." *Annual Review of Political Science* 1(1): 75–93.
- Simmons, Beth A. 2000. "International law and state behavior: Commitment and compliance in international monetary affairs." *American Political Science Review* 94(04): 819–835.
- Simmons, Beth A. 2009. *Mobilizing for Human Rights: International Law in Domestic Politics*. Cambridge: Cambridge University Press.
- Simmons, Beth A., and Lisa Martin. 2012. *International Organisations and Institutions*. Sage Publications pp. 326–351.
- Smith, Jeffrey, and Petra Todd. 2005. "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics* 125(1): 305–353.
- Staiger, Robert W., and Kyle Bagwell. 1999. "An Economic Theory of GATT." *American Economic Review* 89(1): 215–248.
- Stein, Arthur A. 1990. *Why Nations Cooperate: Circumstance and Choice in International Relations*. Ithaca: Cornell University Press.
- Steinberg, Richard H. 2004. "Judicial Lawmaking at the WTO: Discursive, Constitutional, and Political Constraints." *Am. J. Int'l L.* 98: 247.
- Stone Sweet, Alec, and Thomas Brunell. 2012. "The European Court of Justice, state noncompliance, and the politics of override." *American Political Science Review* 106(01): 204–213.
- Subramanian, Arvind, and Shang-Jin Wei. 2007. "The WTO promotes trade, strongly but unevenly." *Journal of International Economics* 72(1): 151 – 175.
- Sweet, Alec Stone, and James A Caporaso. 1998. "From free trade to supranational polity: The European Court and integration." *European integration and supranational governance* pp. 92–133.

- Sweet, Alec Stone, and Wayne Sandholtz. 1997. "European integration and supranational governance." *Journal of European public policy* 4(3): 297–317.
- Tallberg, Jonas. 2000. "Supranational influence in EU enforcement: the ECJ and the principle of state liability." *Journal of European Public Policy* 7(1): 104–121.
- Tallberg, Jonas. 2002. "Delegation to supranational institutions: why, how, and with what consequences?" *West European Politics* 25(1): 23–46.
- Tomz, Michael. 2007. "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61(04): 821–840.
- Tomz, Michael, Joshua A. Tucker, and Jason Wittenberg. 2002. "An Easy and Accurate Regression Model for Multiparty Electoral Data." *Political Analysis* 10(1): 66–83.
- Tsebelis, George. 1995. "Decision Making in Political Systems: Veto Players in Presidentialism, Parliamentarism, Multicameralism and Multipartyism." *British Journal of Political Science* 25(3): 289–325.
- Tsebelis, George. 2003. *Veto Players: How Political Institutions Work*. Princeton: Princeton University Press.
- United Nations. 2013. "Commodity Trade Statistics Database (Comtrade)." *Available from:* <http://comtrade.un.org> .
- Voeten, Erik. 2013. "Does Participation in International Organizations Increase Cooperation?" *The Review of International Organizations* pp. 1–24.
- Von Stein, Jana. 2005. "Do Treaties Constrain or Screen? Selection Bias and Treaty Compliance." *American Political Science Review* 99(4): 611–622.
- World Bank. 2013. "World Development Indicators." *Available from:* <http://data.worldbank.org/indicator> .

- Yarbrough, Beth V, and Robert M Yarbrough. 1987*a*. "Cooperation in the Liberalization of International Trade: After Hegemony, What?" *International Organization* 41(01): 1–26.
- Yarbrough, Beth V., and Robert M. Yarbrough. 1987*b*. "Institutions for the Governance of Opportunism in International Trade." *Journal of Law Economics & Organization* 3: 129.
- Yarbrough, Beth V., and Robert M. Yarbrough. 1992. *Cooperation and Governance in International Trade: The Strategic Organizational Approach*. Princeton: Princeton University Press.