# Domestic Inequality and Exit from International Agreements*

Justin Melnick†

January 30, 2024

**Abstract**

Why do democracies withdraw from international economic agreements? I offer a novel explanation of exit from these treaties: many leaders cannot credibly promise to compensate the losses these agreements cause to voters. I develop a formal model situated in a democracy that is party to a trade agreement that creates "winners" and "losers" in the domestic electorate, and where leaders vary in their ability to implement redistributive policies designed to offset losers' incurred damages. Leaders for whom implementation costs are high often have electoral incentives to propose exit from agreements rather than maintain the globalized status quo, causing withdrawal from agreements to become politically salient. In particular, as the gains from globalization are more unequally distributed between winners and losers, redistribution becomes politically suboptimal, and exit more attractive. Despite globalization being welfare-enhancing, leaders have domestic incentives to withdraw from agreements: it enhances their electoral odds.

Word Count: 7360

*I am grateful to Peter Rosendorff, Sandy Gordon, Alastair Smith, Bruce Bueno de Mesquita, Hye Young You, Matt Malis, Jay Shon, Carolina Torreblanca, Kun Heo, Martin Castillo-Quintana, Alexis Palmer, Roxanne Rahnama, and participants at the NYU IR Workshop, NYU Graduate Political Economy Seminar, NYU Pre-Dissertation Seminar, APSA 2023, and GSIPE for insightful comments and feedback.

†Ph.D. Candidate, Department of Politics, New York University. Contact: melnickj@nyu.edu and justinmelnick.github.io.

Opposition to globalization is increasingly commonplace in advanced democracies. A notable feature of this "globalization backlash" (Walter 2021*a*) is the withdrawal from, halt in negotiations over, or threat to exit from international economic agreements. For example, former President Trump withdrew the U.S. from the Trans-Pacific Partnership, threatened to exit from N.A.F.T.A., and abrogated W.T.O. rules to escalate trade tensions against China, with some in his party calling for American withdrawal from the W.T.O. altogether (Cooley and Nexon 2020). Across the globe, Brexit was in part motivated by fair trade concerns,[1] and Indian Prime Minister Modi withdrew the country from dozens of bilateral investment treaties in 2016.[2] The withdrawal from these types of agreements is puzzling, as globalization is thought to be welfare-enhancing (Cameron 1978; Katzenstein 1985; Rodrik 1997). Why would leaders withdraw from these agreements (von Borzyskowski and Vabulas 2019)? And why would voters reward exit electorally (Flaherty and Rogowski 2021; Milner 2021)?

I build a formal model to explain why leaders exit from international economic agreements grounded in the logic of domestic electoral competition. The model documents a strategic tension between the electoral value of redistribution, the traditional domestic policy instrument used to sustain globalization (Ruggie 1982), and withdrawal from international agreements. Leaders, who vary in the efficiency with which they can implement redistributive policy, may find it politically optimal to withdraw from agreements when globalization's gains are more unequally distributed because redistribution becomes too costly relative to the concomitant electoral returns.

Globalization has increased economic inequality domestically (Rodrik 2018; Catão and Obstfeld 2019), with its benefits accruing to a select group of globalization "winners" at the expense of "losers." While redistribution attempts to address rising inequality by recalibrating relative gains between winners and losers, so too does exit. Leaders thus leverage tensions over domestic inequality politically by proposing to leave international economic agreements. If globalization raises inequality, exit becomes more attractive by empowering globalization's losers with more credible promises of *equality*. I show that, in times where inequality becomes intolerably high, politicians may find it profitable to threaten to de-globalize as a means of securing the electoral support of those harmed by globalization, making exit a salient policy option.

In addition to explaining the ascendance of leaders threatening to exit from agreements, the model also provides basis for another empirical fact, illustrating how rising inequality entices leaders who traditionally supported integration to abandon it in favor of exit, inducing "political realignments"[3] in the composi-

[1] https://www.bbc.com/news/uk-politics-eu-referendum-35783049

[2] https://hsfnotes.com/arbitration/2017/03/16/mixed-messages-to-investors-as-india-quietly-terminates-bilateral-investment-treaties-with-58-countries/

[3] See Schonfeld (2021) and Schonfeld and Winter-Levy (2021) for evidence of realignments on issues of international trade

tion of leaders' bases of electoral support. With low inequality, losers tolerate globalization with requisite compensation, voting for leaders who can efficiently redistribute ("pro-redistribution leaders"). Winners, reluctant to fund compensation (Linardi and Rudra 2020), vote for leaders for whom redistribution is relatively costly ("anti-redistribution leaders"). Concurrently, rising inequality makes leaders more likely to withdraw from international agreements, and I show that anti-redistribution leaders prefer exit more often than pro-redistribution leaders. This creates cases in which exit is electorally salient: the pro-redistribution leader upholds globalization but the anti-redistribution leader proposes exit. Voters subsequently shift their political allegiances. Winners, who stand to lose substantial gains if exit occurs, now support the pro-redistribution leader despite having to finance compensation. Losers, no longer viewing such redistribution as sufficient compensation (Bowen, Broz and Rosendorff 2022), support leaders who would exit instead.

Importantly, by microfounding the logic of political realignments, the model shows that parties supporting globalist endeavors need not remain the same over time: leaders' support of globalization policy depends on the extant level of inequality in society and how they can expect to capture different blocs of the electorate. Scholars have extensively interrogated the link between anti-globalization sentiment and support for far right parties (e.g., Colantone and Stanig 2018; Autor et al. 2020). However, candidates on the left have also supported withdrawal from international agreements because of their consequences for domestic inequality: the model demonstrates rationalizes this when inequality is exceedingly high. This paper thus elucidates how concerns for domestic political survival trump support for globalization because leaders have incentives to sacrifice welfare-enhancing liberalization for electoral expedience.

## Contribution

This paper's contribution is twofold. First, I introduce a simple theoretical model to explain how domestic political competition incentivizes withdrawal from international agreements and makes the threat of exit an electorally salient issue. To date, this analysis provides one of the first stories to rationalize leaders' decisions to withdraw from international agreements as a function of their domestic politics. Moreover, since conventional economic wisdom would hold that an optimal transfer from winners to losers should always exist, thus obviating the possibility of exit, it is imperative to elucidate the causal mechanism through which domestic politics shapes the feasibility and salience of anti-globalist policy.

Two recent empirical papers highlight the domestic political tensions that the model describes. Flaherty and Rogowski (2021) demonstrate that "top-heavy inequality," a distribution of earnings concentrated within

and European integration.

a small, "elite" faction of society, conditions support for anti-globalist or populist leaders. They document that rising inequality is necessary to elicit voter support for anti-integrationist leaders. Milner (2021) shows that increased exposure to trade increases support for extreme right parties, and that social welfare programs appear not to dampen or reverse trends of far-right voting. Together, this work suggests that rising inequality due to globalization precipitates anti-integration preferences and that compensation fails to moderate these preferences. This presents a clear opportunity for theoretical work to clarify the underlying causal mechanism.

The paper's second contribution relates specifically to the study of exit as an outcome, where I depart from extant "state-level" arguments. Three prevailing arguments exist to explain why states withdraw from international agreements, all of which treat nations as black boxes. Most prominent is a story about "composition effects," which argues that preference divergence among member states over time leads to withdrawal from agreements (von Borzyskowski and Vabulas 2019; Malis, Rosendorff and Smith 2022). Scholars point to empirical differences in regime type or changes in ideal points across member states to justify a country's exit from an agreement. Yet, while it may be the case that withdrawing governments over time have become disgruntled with IO performance (von Borzyskowski and Vabulas 2019), ostensibly lending credence to a story based on composition effects, any "changes in state preferences" are endogenous to domestic political changes. Increased incidence of exit may also be due to contagion effects (e.g., Walter 2021*b*), whereby withdrawal by one state motivates others to follow suit, similar to a logic of unraveling from a previously-established cooperative equilibrium (Schelling 1960). However, advocates of contagion fail to identify the motives of the "first mover," or why one state exits in the first place. A final strand of literature has cast exit as a consequence of growing regime complexity and bounded rationality, in which exit from some international commitments becomes inevitable when they are superseded by less constricting forms of integration (Haftel and Thompson 2018). My explanation for the emergence of leaders promoting exit rests on a "two-level game" logic in which leaders interface with international commitments subject to domestic constraints, and international conditions shape domestic political outcomes (cf. Putnam 1988; Milner and Rosendorff 1997; Buisseret and Bernhardt 2018; Melnick and Smith 2023).

## Empirical Motivation

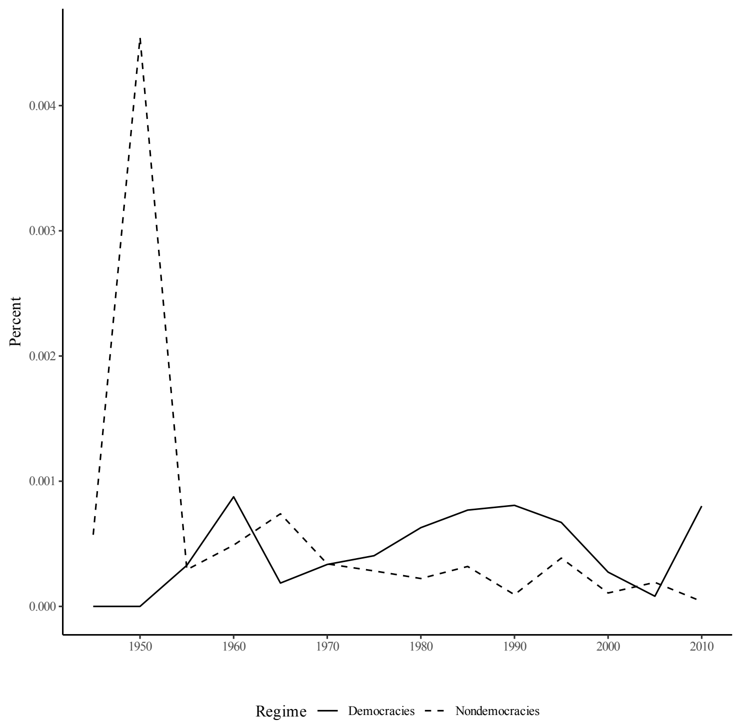
To motivate the model, consider some empirical description about countries' withdrawals from intergovernmental organizations over time. I use data from von Borzyskowski and Vabulas (2019), the most comprehensive empirical study of exit from international organizations, who describe withdrawals from intergovernmental organizations between 1945 and 2014 as documented by the Correlates of War Project (Pevehouse et al. 2020). Between 1945 and 2014, there were 188 exits from intergovernmental organizations. As the number of recorded organizations has increased since 1945, so have withdrawals from around 2 organizations per year before 1980 to about 3.5 organizations per year after 1980 (von Borzyskowski and Vabulas 2019, 340). Note that the Correlates of War Project records data on intergovernmental organizations, which only includes organizations that have at least three members, hold regular plenary sessions at least once every ten years, and possess a permanent secretariat and corresponding headquarters; this data is therefore a lower bound on the number of observed exits.

I augment these analyses by considering the regime type of member-states that withdraw from intergovernmental organizations and the type of organizations from which these states withdraw. The top panel of Figure 1 depicts the number of withdrawals from intergovernmental organizations by democracies (solid line) and nondemocracies (dashed line), normalized by both the number of democracies/nondemocracies and number of intergovernmental organizations in force in a given year.[4] In the years following the Second World War, there was a spike of nondemocratic states exiting from agreements, namely withdrawals from the World Health Organization by Albania, Bulgaria, Czechoslovakia, Poland, and Romania in 1950. After 1950, democratic states seem to exit from intergovernmental organizations at a greater frequency than nondemocratic states, and consistently do so after 1970. Indeed, as von Borzyskowski and Vabulas (2019, 342) document, the top exiters during this time period are the United States (10), Canada, (9), the United Kingdom (7).

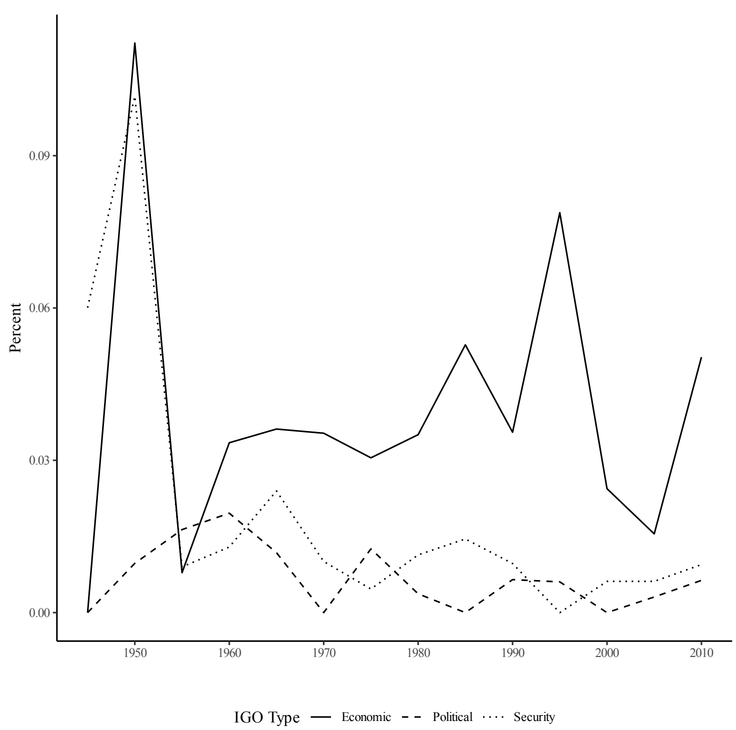
Beginning around 2005, withdrawals by democratic states appear to be on the rise. This trend comports with more contemporary cases like Brexit in 2016, Israel's exit from UNESCO in 2019, and the United States' withdrawal from the International Coffee Organization in 2018 and UNESCO in 2019. Understanding the trend of democratic exits in an era of hyperglobalization warrants a theoretical model that can provide leverage for the causal mechanism at play.

I also consider the types of agreements from which states exit. This is illustrated in the bottom panel of Figure 1, which plots the number of exits from intergovernmental organizations classified as economic agreements (solid line), political agreements (dashed line), or security agreements (dotted line), normalized by the number of functional intergovernmental organizations in a given year. In total, 127 withdrawals were from organizations classified as economic, while 42 withdrawals were from security agreements (including the World Health Organization), and 19 were withdrawals from political organizations. Examples of withdrawals from

[4] I define a democracy as a state that scores at least 7 on the Polity V measure (Marshall and Gurr 2020).




(a) Percent of IO Withdrawals by Regime Type




(b) Percent of IO Withdrawals by Issue Area

Figure 1: IO Withdrawals by Regime Type and Issue Area, 1945-2014

economic organizations include Mauritania's exit from the West African Monetary Union in 1973, Canada's exit from the Inter-American Tropical Tuna Commission in 1984, and Norway's exit from the International Copper Study Group in 2000. The percentage of exits from economic agreements is consistently larger than the percentage of exits from other types of agreements; for this reason I scope the model to thinking about why states may wish to exit from international economic agreements.

These two stylized facts – that democracies seem to be exiting from intergovernmental organizations more frequently than nondemocracies and that states seem to be exiting from economic organizations – motivate the model. Evidently, the data demonstrate that exiting from international agreements is a rare event. Indeed, as von Borzyskowski and Vabulas (2019) acknowledge, states often threaten to exit from international agreements much more than they actually withdraw. This underscores the need to understand how exit – or the threat of exit by domestic political actors – can emerge in equilibrium.

# Model

I study a domestic political contest within a democratic country party to an international economic agreement.[5] Two possible leaders, $L$ (she) and $H$ (he), advance policy proposals to attract the votes of a continuum of individuals indexed by $i \in [0, 1]$. Proposals jointly determine how much redistribution should occur within society and whether the country remains in or exits from the agreement. I assume that these proposals are made simultaneously and bind in a Downsian fashion to study how electoral incentives affect the likelihood of remaining in or exiting from the agreement.

Globalization has three distinct effects. First, it produces aggregate welfare gains. In the agreement, the economy is of size $\gamma > 1$. Second, the agreement fosters "winners" and "losers" domestically. Let some share $\lambda \in [0, 1]$ of voters be globalization losers ($\ell$) while a share $1 - \lambda$ are globalization winners ($w$). In reduced form, a fraction $1 - \lambda$ of society is employed in sectors that benefit from the agreement, while a share $\lambda$ is adversely affected. Third, globalization determines the share of the economy that each of the two groups commands. Under globalization, the status quo, winners own a share $\alpha_R \in [0, 1]$ of the economy and losers own $1 - \alpha_R$. The parameter $\alpha_R$ represents the structural factors that affect returns to income for winners and losers in the agreement. Thus, globalization determines aggregate gains ($\gamma$), who benefits or does not benefit from these gains ($\lambda$), and the relative distribution of gains ($\alpha_R$).

The $1 - \lambda$ individuals are "winners" because their *ex-ante* per capita income in the agreement is greater

[5] The setup most closely resembles a reduced-form modeling of an international trade agreement, but any such economic agreement with distributional conflict would also fit within the model's scope.

than the losers'. Globalization makes some domestic groups better off than others. I therefore restrict attention to the case where $\frac{\alpha_R}{1-\lambda} > \frac{1-\alpha_R}{\lambda}$ or $\alpha_R > 1-\lambda$. This condition ensures that, absent any redistribution, winners are better off than losers when the agreement is in place.

By exiting the agreement, the economy contracts to size 1. The overall contraction of the economy represents aggregate welfare losses when gains from trade dissipate; viz. the costs of Brexit on the U.K. economy have been estimated to be about \$124bn per year in lost output.[6] Note that while exit may be costly in other senses besides shrinking overall welfare – perhaps leaders who exit or threaten to exit from agreements suffer reputational costs, etc. – I do not model them without loss of generality. Such costs are orthogonal to the mechanism under study, and as will be shown, any cost levied upon leaders that is not leader-specific leaves the results unchanged.

In addition to aggregate losses, winners' relative gains contract to $\alpha_E \leq \alpha_R$. This assumption reflects the idea that winners lose some of their gains from specialization if exit occurs, or that, prior to any government intervention, income accruing to winners is greater when the country is party to the agreement. Moreover, this assumption encodes the idea that globalization-related inequality and aggregate gains from globalization erode together (Antràs, de Gortari and Itskhoki 2017). Note further that $\frac{\alpha_R}{1-\alpha_R}$, the income ratio under globalization, can be thought of as a measure of inequality between winners and losers. Increasing the *ex-ante* gains to winners $\alpha_R$ creates more inequality. I will use "inequality" and "winners' gains from globalization" to refer to $\alpha_R$ interchangeably. By definition, exiting from the agreement necessarily makes society more equal, $\frac{\alpha_E}{1-\alpha_E} \leq \frac{\alpha_R}{1-\alpha_R}$.[7]

Each leader's proposal can be written as a share of national income to be allocated to winners, $\theta_{da}$ for $d \in \{L, H\}$ and $a \in \{remain, exit\}$. After $L$ and $H$ have announced these proposals, citizens go to the polls. Voters have increasing, concave payoffs $v(x) = \log(x)$ over income. This means that an individual benefits more when their group enjoys a greater share of national income, but exhibit diminishing marginal returns. The following table summarizes the per capita income to winners and losers for each leader's proposals under remain and exit.

Voting is based on a standard probabilistic voting model (Lindbeck and Weibull 1987). Individuals vote sincerely by comparing the differences in their expected income under $L$ versus $H$, also taking into account individual-specific shocks $\mu_i$ and a common preference shock $\beta$ "in favor" of $L$. The shocks represent voters'

[6] https://www.bloomberg.com/news/articles/2023-01-31/brexit-is-costing-the-uk-100-billion-a-year-in-lost-output?in_source=embedded-checkout-banner.

[7] If the difference $\alpha_R - \alpha_E$ is small, then so too will be the change in inequality induced by exit, which may reflect a setting in which countries are more tightly integrated with one another so the marginal effect of exiting one agreement may be smaller on inequality. By contrast, if $\alpha_R - \alpha_E$ is large, then exit has more consequential effects on globalization-related inequality. I do not require $\alpha_R - \alpha_E$ to be large for any of the results.

| | $d = L$ | | $d = H$ | |
|---|---|---|---|---|
| | Remain | Exit | Remain | Exit |
| Winners $w$ | $\frac{\gamma\theta_{LR}}{1-\lambda}$ | $\frac{\theta_{LE}}{1-\lambda}$ | $\frac{\gamma\theta_{HR}}{1-\lambda}$ | $\frac{\theta_{HE}}{1-\lambda}$ |
| Losers $\ell$ | $\frac{\gamma(1-\theta_{LR})}{\lambda}$ | $\frac{1-\theta_{LE}}{\lambda}$ | $\frac{\gamma(1-\theta_{HR})}{\lambda}$ | $\frac{1-\theta_{HE}}{\lambda}$ |

Table 1: Income Distribution across Winners and Losers

valuations of $L$ on all other electorally relevant issues besides globalization policy. Let $\mu_i \sim U[-\frac{1}{2m}, \frac{1}{2m}]$ and $\beta \sim U[-\frac{1}{2b}, \frac{1}{2b}]$. The parameters $m$ and $b$ define the salience of globalization policy relative to other issues in the electoral landscape. When $m$ and $b$ are large, globalization policy looms heavily on voters' assessments of political leaders. Conversely, when $m$ and $b$ are small, leaders' globalization policy proposals carry little weight in voting decisions.

Let $v_{ij}(\theta_{da})$ be individual $i$ in group $j$'s utility from the proposal $\theta_{da}$, as described in Table 1. Voter utility is therefore

$$u_{ij} = \begin{cases} v_{ij}(\theta_{La}) + \mu_i + \beta & \text{vote for } L. \\ v_{ij}(\theta_{Ha}) & \text{vote for } H. \end{cases}$$

Leaders consider how enacting their redistributive proposals affects the chances of winning the election as well as the costs of implementing these proposals. Domestic redistribution has two costs in the model. The first is electoral: any redistribution toward losers might buy their votes, albeit at the expense of support from the winners. Increasing the wellbeing of one group necessarily diminishes political support from the other. The second cost concerns the implementation of redistributive policy. I assume it is costly to move policy away from the status quo $\alpha_R$, operationalized as a marginal cost of policymaking $\kappa_d$. Some leaders can redistribute more cheaply than others. $L$'s implementation costs are "low," $\kappa_L = 1$, while $H$'s are "high," $\kappa_H = \kappa > 1$. These costs could represent intrinsic leader traits – like their managerial competence or their exogenous ideological predisposition toward domestic redistribution (Boix 1998) – or literal costs undertaken by a leader's administration to execute redistributive policy, which could be microfounded within the bureaucracy. One may expect leaders from different parties to appoint bureaucrats with distinct policy preferences over redistribution (Halperin and Clapp 2006); higher implementation costs could stem from bureaucratic resistance to enacting redistributive policy.

Leaders care exclusively about their electoral prospects and have no innate preferences over how income should be distributed in society. When proposing policies, $L$ and $H$ maximize their chances of winning the election less redistribution's implementation costs. If elected, leaders enjoy a benefit $\Psi > 0$ and receive a payoff of zero if they lose the election. If $\pi(\theta_{La}, \theta_{Ha})$ is the (endogenous) probability that $L$ wins the election given proposals $\theta_{da}$, then leaders choose $\theta_{da}$ to maximize

$$EU_L(\theta_{La}, \theta_{Ha}) = \pi(\theta_{La}, \theta_{Ha})\Psi - \frac{1}{2}(\alpha_a - \theta_{La})^2.$$

$$EU_H(\theta_{La}, \theta_{Ha}) = (1 - \pi(\theta_{La}, \theta_{Ha}))\Psi - \frac{\kappa}{2}(\alpha_a - \theta_{Ha})^2.$$

Based on the optimal $\theta_{da}$s, leaders subsequently determine whether to remain in or exit from the agreement.

A strategy for each leader is to propose whether to remain in or exit from the agreement, and how to divide the pie in each of these two eventualities. A strategy for voter $i$ is a choice to vote for $L$ or $H$ given platform proposals and draws of the shocks. I solve for the subgame perfect equilibrium of the game via backward induction. All proofs are in the appendix. Table 2 summarizes model notation.

| | |
|---|---|
| $j \in \{w,\ \ell\}$ | Subscript for if individual is a globalization winner or loser |
| $d \in \{L,\ H\}$ | Subscript for leader $L$ or $H$ |
| $a \in \{R,\ E\}$ | Subscript for international outcome, remaining in or exiting from the agreement |
| $\gamma > 1$ | Size of the globalized economy |
| $\lambda \in [0, 1]$ | Share of globalization losers |
| $\alpha_a \in [0, 1]$ | Status quo distribution of income to winners |
| $\theta_{da} \in [0, 1]$ | Redistributive proposal of leader $d$ |
| $v_{ij}(x)$ | Individual utility over income |
| $\mu_i \sim U[-\frac{1}{2m},\ \frac{1}{2m}]$ | Voter-specific valence shock |
| $\beta \sim U[-\frac{1}{2b},\ \frac{1}{2b}]$ | Common valence shock |
| $\kappa_L = 1$, $\kappa_H > 1$ | Leader's marginal costs of implementing redistributive policy |
| $\Psi > 0$ | Leader's benefit to winning the election |
| $\pi(\theta_{La}, \theta_{Ha}) \in [0, 1]$ | Probability that $L$ wins election given proposals $(\theta_{La}, \theta_{Ha})$ |
| $D_j = v_{ij}(\theta_{La}) - v_{ij}(\theta_{Ha})$ | Difference in utility for group $j$ given proposals $(\theta_{La}, \theta_{Ha})$ |

Table 2: Model Notation

# Analysis

As in other probabilistic voting models, comparing voters' expected utilities from voting for each leader yields an endogenous probability that $L$ wins the election $\pi(\theta_{La}, \theta_{Ha})$ given any pair of redistributive proposals $(\theta_{La}, \theta_{Ha})$. There are four scenarios that voters face. Both leaders could propose to remain in the agreement, both could propose to exit, and one could propose remain while the other proposes exit. To determine the optimal retention rule, voters prospectively evaluate their differences in expected income between $L$ and $H$. This is a simple comparison of the utility proposed by each leader, as in Table 1. Define $D_j$ as the difference in income that a member of group $j$ would incur when voting for $L$ versus $H$. Voter $i$ in group $j$ votes for

$L$ whenever $v_{ij}(\theta_{La}) + \mu_{ij} + \beta \geq v_{ij}(\theta_{Ha})$, or when

$$\underbrace{v_{ij}(\theta_{La}) - v_{ij}(\theta_{Ha})}_{D_j} + \mu_{ij} + \beta \geq 0.$$

Then, for any values of $D_w$ and $D_\ell$, I obtain a general expression for the probability that $L$ wins the election. Leaders' choices of globalization policy factor into the chances of electoral success through a simple population-weighted average of the differences in voters' expected income, which is written as

$$\pi\Big(D_w(\theta_{La}, \theta_{Ha}), D_\ell(\theta_{La}, \theta_{Ha})\Big) = \frac{1}{2} + b\Big((1-\lambda)D_w + \lambda D_\ell\Big).$$

## Redistribution and Exit in Equilibrium

I focus on leaders' motivations to redistribute and whether to exit from the agreement. In equilibrium, leaders choose $\theta^*_{da}$ to balance the marginal electoral benefits with the marginal costs of implementing redistribution.

**Proposition 1** *The equilibrium proposal by leader $d$ in outcome $a$ solves*

$$b\Psi\frac{\lambda}{1-\theta^*_{da}} = b\Psi\frac{1-\lambda}{\theta^*_{da}} + (\alpha_a - \theta^*_{da})\kappa_d.$$

Leaders find the optimal transfer of income from winners to losers that is electorally maximizing. The left-hand side of the equation in Proposition 1 represents the marginal electoral benefits of redistribution by providing a greater share of income to losers. The right-hand side is the marginal opportunity cost of redistribution, which encapsulates both the decrease in electoral support from winners as well as implementation costs. Redistribution is politically costly and administratively costly. Since she can redistribute more cheaply, the low-implementation cost leader $L$ will always propose greater redistribution than the high-cost leader $H$.

**Corollary 1** *$L$ redistributes more than $H$, $\theta^*_{La} \leq \theta^*_{Ha}$.*

In a fairly straightforward way, $L$ can redistribute more than $H$ because her marginal costs of implementing redistributive policy are smaller than $H$'s, $\kappa_H > 1 = \kappa_L$. Note, however, that without implementation costs ($\kappa_H, \kappa_L \to 0$), redistribution would be proportional to group size, $\theta^*_{da} = 1 - \lambda$ and would not vary across political leaders. However, the introduction of implementation costs means that *ex-ante* inequality matters for redistribution: $\theta^*_{da} > 1 - \lambda$ because leaders face an additional cost from moving policy away from the status quo. Consequently, increasing inequality constrains the amount that leaders are willing to redistribute. While there are electoral returns to redistribution, it becomes less attractive to move policy away from the status quo as $\alpha_R$ increases. Counterintuitively, leaders must promise more to the winners when inequality is high. In a hyperglobalized world where the distribution of gains is highly unequal, compensation to losers decreases precisely when it is needed most.

**Corollary 2** *Equilibrium proposals $\theta^*_{dR}$ are increasing in the winners' gains $\alpha_R$.*

Given the optimal proposals $\theta^*_{da}$, leaders determine whether it is politically profitable to remain in the agreement or to exit. Leaders exit if and only if globalization's aggregate gains are not too large.

**Proposition 2** *There exist thresholds $\gamma_L \leq \gamma_H$ such that leader $d$ exits the agreement whenever $\gamma \leq \gamma_d$ and remains otherwise.*

Globalization is welfare-enhancing and exit shrinks the size of the pie. If the aggregate gains from globalization $\gamma$ are large, then exit cannot occur. However, if $\gamma$ is relatively meagre, then leaders may find it politically opportunistic to forsake a larger pie for the possible electoral benefits associated with exiting from the agreement, winning the support of the losers. In this case, the "second-order" domestic distributive frictions that globalization creates outweigh the overall gains (Rodrik 2018).

Exit is thus a purely political endeavor that arises because leaders differ in their abilities to implement redistributive policy. Without implementation costs, leaders would never exit from the agreement. Leaders with higher costs of redistribution (higher $\kappa_d$) propose smaller transfers from winners to losers and are more likely to exit from agreements. High-cost $H$ is less able to make the requisite transfers under a globalized regime, and would rather abrogate the agreement. By contrast, since $L$ can redistribute more cheaply, she always has greater incentives to uphold liberalization through redistribution. Therefore, for intermediate values of $\gamma$, the issue of whether to remain in or withdraw from the agreement becomes electorally salient because $L$ would remain but $H$ would exit: domestic political competition determines the fate of globalization.

Exit necessarily reduces inequality between winners and losers because winners lose some of their gains when the agreement is abrogated as the income distribution shifts from $\alpha_R$ to $\alpha_E$. This functions as if redistribution were costless as the economy recalibrates to a new status quo. Despite destroying surplus, exit can be politically enticing because of its appeal of equality toward globalization's losers, especially in the wake of rising inequality. As such, leaders become more likely to exit from international agreements when inequality rises.

**Corollary 3** *The thresholds $\gamma_d$ are increasing in the gains to winners $\alpha_R$.*

When the gains from globalization are unequally distributed domestically, redistribution becomes politically suboptimal because the costs of implementing redistributive policy impede the disbursement of sufficient transfers. Consequently, exit becomes more attractive because rejecting international integration serves as a means of generating a more equal income distribution domestically at a lower cost to leaders. These distributive concerns become more prevalent when $\alpha_R$ increases and compensation is more difficult to enact: globalization's "first-order" welfare gains must increase concomitantly for the agreement to survive if domestic inequality becomes more severe. The connection between inequality and the proposal of exit is illustrated in Figure 2, which plots the thresholds $\gamma_L$ (solid line) and $\gamma_H$ (dotted line) as a function of *ex-ante* inequality under the agreement $\frac{\alpha_R}{1-\alpha_R}$, which are increasing in $\alpha_R$. To the right of each line represents regions of the parameter space where leaders would be willing to exit the agreement.

## Political Realignments

I now use the model to explain another empirical regularity in the era of the globalization backlash: the ascendance of leaders in traditionally pro-globalization parties supporting anti-globalization policies like exit, and the embrace of pro-globalization policies by traditionally more protectionist parties (Schonfeld 2021; Bowen, Broz and Rosendorff 2022). I call these shifts "political realignments" and demonstrate how they depend on rising globalization-induced inequality.

Rising inequality evidently alters leaders' valuations of globalization and subsequently affects which types of voters support them electorally. All else equal, voters prefer the policy proposal that provides their group with the highest per capita income. Since leaders' redistributive proposals and willingness to support globalization are sensitive to rising inequality, so too is voter behavior.

Consider the case where the size of the pie is so large that the globalized status quo would always persist, $\gamma > \gamma_H$. Since liberalization with redistribution is the only credible outcome, losers support the pro-redistribution $L$ who can compensate them more easily, and winners support $H$. However, if inequality rises such that one domestic leader finds it optimal to exit, political alliances reverse. Formally, increasing $\alpha_R$ moves the cutpoints $\gamma_d$ such that, for a fixed $\gamma$, domestic political conditions now motivate $H$ to exit the agreement while $L$ would remain, $\gamma_L \leq \gamma \leq \gamma_H$. $H$'s ability to withdraw from the agreement under

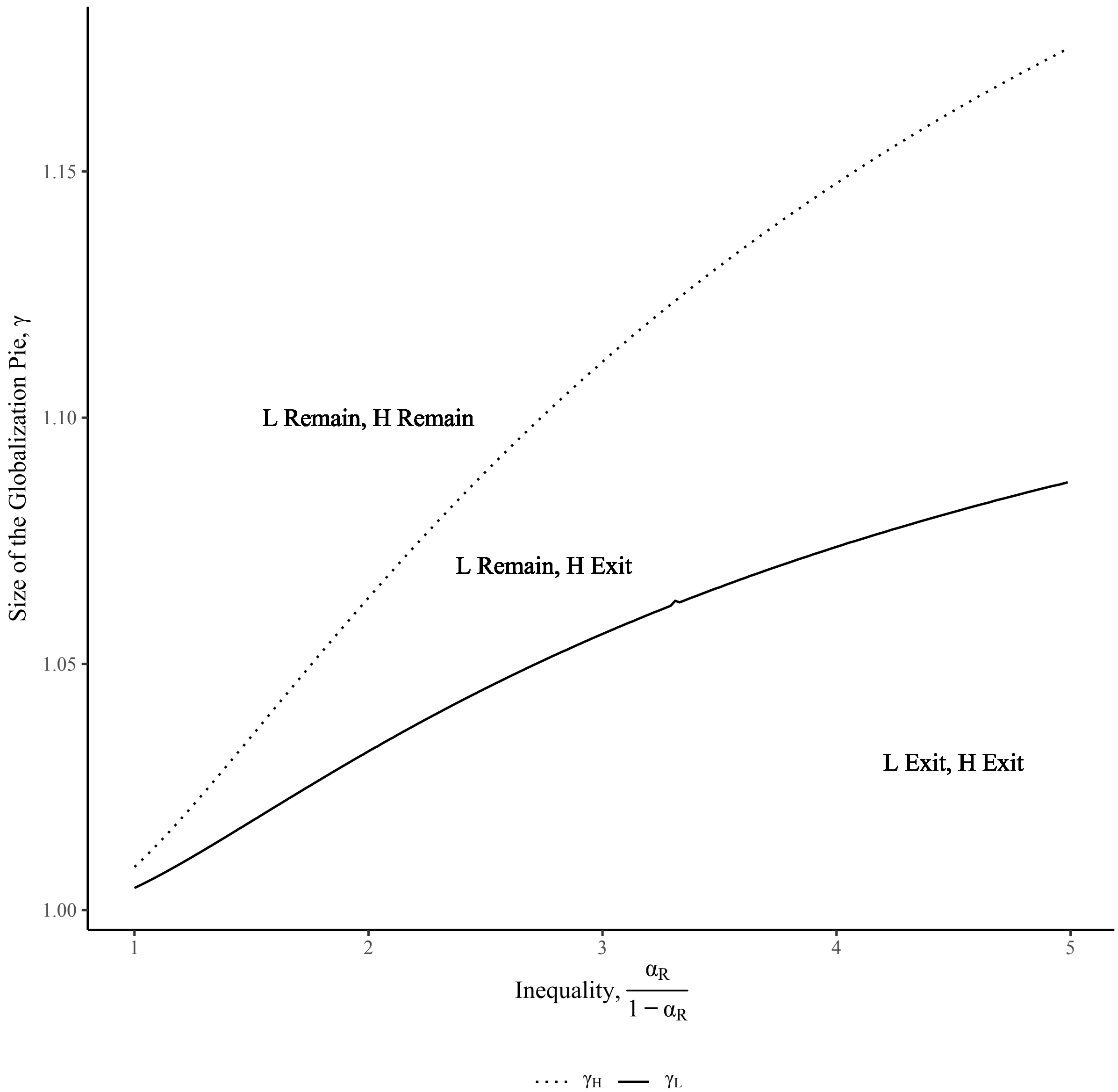



Figure 2: Inequality and Leaders' Thresholds for Exit

conditions where $L$ cannot delivers him an entirely new group of political supporters, the globalization losers, by promising greater equality through exit than with compensation. Rather than supporting the pro-redistribution $L$, losers back $H$, the anti-redistribution, anti-globalization leader. This is because for losers, exit delivers a larger share of a smaller pie, $1 - \theta^*_{HE} \geq \gamma(1 - \theta^*_{LR})$. Conversely, winners flock to $L$ who, despite proposing greater redistribution, allocates winners a larger share of the larger, globalized pie, $\gamma\theta^*_{LR} \geq \theta^*_{HE}$. Voters thus undergo a realignment in political support as inequality shifts leaders' optimal globalization policies.

**Proposition 3** *Increasing the winners' gains from globalization $\alpha_R$ can create political realignments.*

Political realignments imply differences in levels of political support as a result of rising inequality. Let $D^*_j$ be the difference in income for an individual in group $j$ voting for $L$ versus $H$ given leaders' equilibrium policy proposals. Any value $D^*_j > 0$ implies that an individual in group $j$ would vote for $L$ in expectation, while $D^*_j < 0$ is an expected vote for $H$. For example, for any size of the pie $\gamma$ in which both leaders support globalization in equilibrium, $\gamma > \gamma_H$, losers in expectation support $L$, $D^*_\ell > 0$ and winners support $H$, $D^*_w < 0$. Changing the distribution of winners' gains $\alpha_R$ shifts the conditions under which exit becomes optimal, holding $\gamma$ constant; when "moving" into an equilibrium where $H$ exits but $L$ does not, $\gamma_L \leq \gamma \leq \gamma_H$, it is now the case that $D^*_\ell < 0$ and $D^*_w > 0$, documenting shifts in political support.

Political realignments can be seen in Figure 3, which plots $D^*_w$ (solid line) and $D^*_\ell$ (dashed line), the differences in income that winners and losers expect when voting for $L$ versus $H$ as a function of *ex-ante* inequality in the agreement. Each segment represents how $D^*_w$ and $D^*_\ell$ change conditional on the policy outcomes that the leaders propose. Consider how the sign of $D^*_j$ changes at the discontinuities between the segments. These discontinuities represent leaders shifting their support away from integration and toward exit.

In the leftmost case, winners' gains $\alpha_R$ are low and the aggregate gains from globalization are high; both leaders would remain in the agreement. Political coalitions arise from the fact that $H$'s willingness to redistribute is limited, leaving winners with greater income with $H$ in power rather than $L$. $L$'s support comes from losers, to whom she can more cheaply redistribute. At the first discontinuity, inequality has grown such that $H$'s optimal action is now to exit rather than remain, which induces a realignment of political support. Now, winners coalesce behind $L$ because they incur a large opportunity cost of abandoning the globalized status quo. By contrast, losers rally behind $H$, who, despite a welfare-inferior foreign policy, ultimately promises them greater income than $L$ would. Finally, when $\alpha_R$ reflects an incredibly unequal

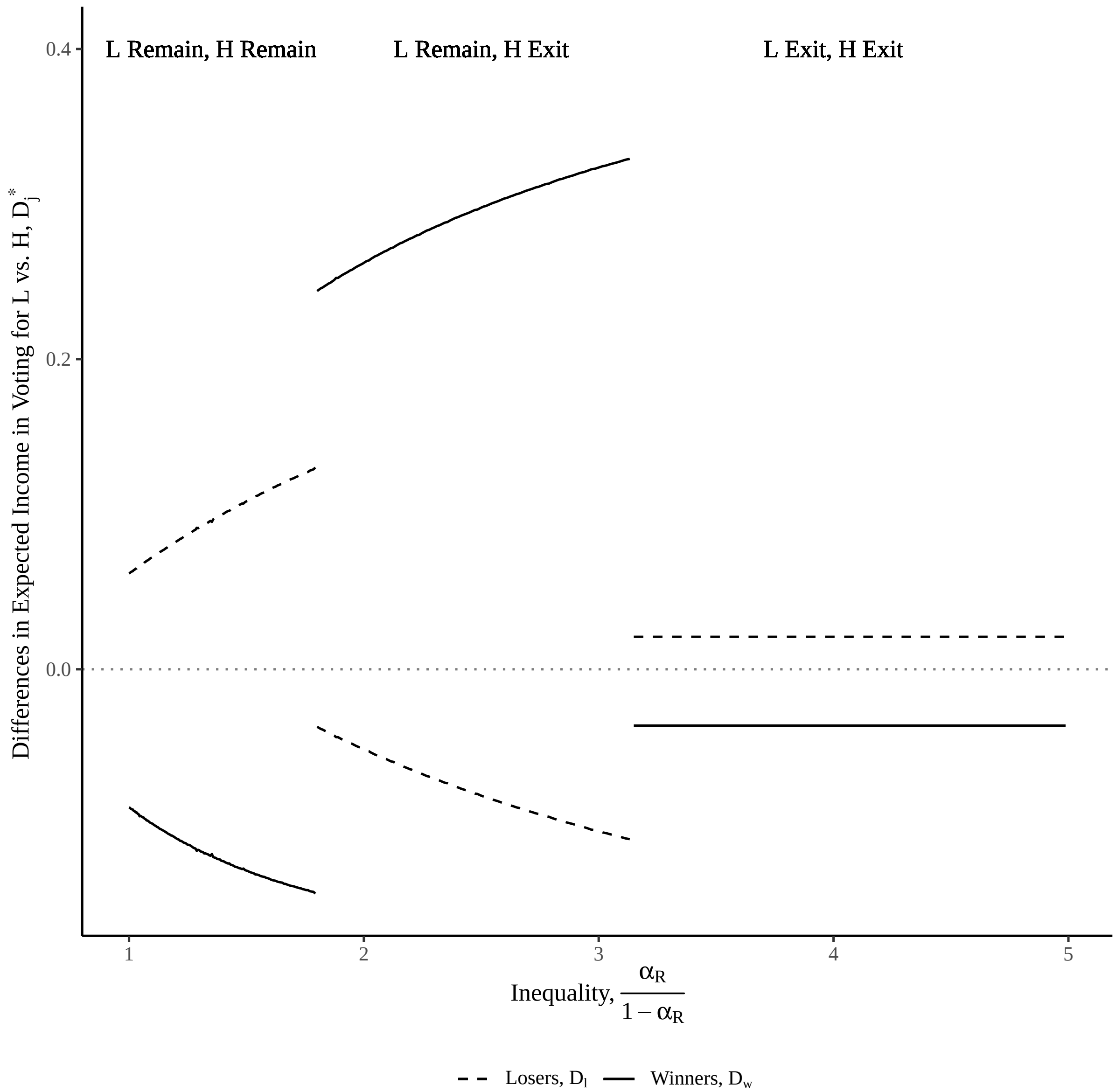



Figure 3: Inequality and Political Realignments

distribution of gains between winners and losers, a second discontinuity occurs. $L$ too would now withdraw from the agreement, and political support switches again. Despite being empirically unlikely, this outcome may characterize a scenario in which inequality is so extreme that it is no longer politically feasible to uphold the agreement. Similar to the first case, winners support $H$ because he can promise them a greater share of the new status quo distribution, $\theta^*_{HE} \geq \theta^*_{LE}$. Losers support $L$ because she can apportion them a greater share of an albeit smaller, de-globalized pie.

Understanding political realignments can help to explain the variation in policy support and subsequent composition of the winning coalition of the Republican Party in the United States. Since the mid-twentieth century, Republicans traditionally supported globalization winners and elites, preferring fewer tariffs and lower levels of redistribution than Democrats (Irwin 2017). Republicans upheld the United States' commitment to international integration, reinforcing support from the globalization winners while simultaneously advancing domestic policies to cut rather than expand redistribution to the globalization losers. Indeed, localities exposed most heavily to globalization saw some of the greatest decline in compensation (Autor, Dorn and Hanson 2013).

Growing inequality in the U.S. due to globalization has shifted the willingness of Republican political leaders to continue to support pro-globalization policies. The rise of right-wing, anti-globalization politicians even predates the presidency of Donald Trump (Cerrato, Ferrara and Ruggieri 2018; Kuk, Seligsohn and Zhang 2022). As early as 1992, Pat Buchanan repudiated support for N.A.F.T.A., and won 23% of the vote in the Republican presidential primary.[8] More recently, Senator Josh Hawley introduced a joint resolution to withdraw the U.S. from the W.T.O. in 2020.[9] These politicians found it electorally advantageous to abandon their support of globalization, and consequently the locus of their political support transitioned from globalization winners to losers, with a majority of Republican voters viewing free trade as a "bad thing" for the U.S. economy by the 2016 election.[10] Moreover, Democratic willingness to support globalization endeavors surged with President Clinton's ratification of N.A.F.T.A. and is further exemplified by the extended support for free trade in the Obama years by concluding free trade agreements with South Korea, Colombia, and Panama, demonstrating partisan reversal in support for globalization policy.

[8] https://www.washingtonpost.com/archive/opinions/1993/11/07/america-first-nafta-never/c8450c08-b14b-4a25-abe8-0b7cfc992e11/.

[9] https://www.hawley.senate.gov/sites/default/files/2020-05/Hawley-WTO-Resolution.pdf.

[10] https://www.pewresearch.org/short-reads/2016/03/31/republicans-especially-trump-supporters-see-free-trade-deals-as-bad-for-u-s/.

# Conclusion

This paper proposes a formal model that explains how rising inequality can lead political leaders to support withdrawal from international agreements. The model points to the electoral incentives that leaders may have to threaten to de-globalize, and elucidates the conditions under which the possibility of withdrawal may be politically salient, namely when globalization-related inequality is on the rise. To be clear, I do not purport to "predict" exit empirically, as anti-globalization candidates can lose the election in the model, but rather to provide theoretical structure on the domestic political incentives that may behoove leaders to support measures like exit that are welfare-inferior. Indeed, observing exit is a rare event (von Borzyskowski and Vabulas 2019) but the model can explain the recent rise in anti-globalization sentiment stemming from various political leaders.

This study helps us to make sense of how political behavior and globalization policy evolve with rising inequality and betters our understanding of how political leaders integrate domestic and foreign policy to advance their electoral objectives. Yet, further research may investigate the mechanisms that incentivize leaders in nondemocratic states to withdraw from international agreements (e.g., Rudra, Nooruddin and Bonifai 2021). Additionally, scholars may wish to interrogate whether withdrawals from other types of organizations beyond economic agreements also have domestic microfoundations.

# References

Antràs, Pol, Alonso de Gortari and Oleg Itskhoki. 2017. "Globalization, inequality and welfare." *Journal of International Economics* 108:387–412.

Autor, David, David Dorn, Gordon Hanson and Kaveh Majlesi. 2020. "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure." *American Economic Review* 110(10):3139–3183.

Autor, David H., David Dorn and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review* 103(6):2121–2168.

Boix, Carles. 1998. *Political Parties, Growth and Equality: Conservative and Social Democratic Economic Strategies in the World Economy.* Cambridge: Cambridge University Press.

Bowen, T. Renee, J. Lawrence Broz and B. Peter Rosendorff. 2022. "Trade Policy Transitions: Three Eras of U.S. Trade Policy.".

Buisseret, Peter and Dan Bernhardt. 2018. "Reelection and Renegotiation: International Agreements in the Shadow of the Polls." *American Political Science Review* 112(4):1016–1035.

Cameron, David R. 1978. "The Expansion of the Public Economy: A Comparative Analysis." *The American Political Science Review* 72(4):1243–1261.

Catão, Luís and Maurice Obstfeld. 2019. *Meeting Globalization's Challenges: Policies to Make Trade Work for All.* Princeton: Princeton University Press.

Cerrato, Andrea, Federico Maria Ferrara and Francesco Ruggieri. 2018. "Why Does Import Competition Favor Republicans?".

Colantone, Italo and Piero Stanig. 2018. "The Trade Origins of Economic Nationalism: Import Competition and Voting Behavior in Western Europe." *American Journal of Political Science* 62(4):936–953.

Cooley, Alexander and Daniel Nexon. 2020. *Exit from Hegemony: The Unraveling of the American Global Order.* New York: Oxford University Press.

Flaherty, Thomas M. and Ronald Rogowski. 2021. "Rising Inequality As a Threat to the Liberal International Order." *International Organization* 75(2):495–523.

Haftel, Yoram Z. and Alexander Thompson. 2018. "When do states renegotiate investment agreements? The impact of arbitration." *The Review of International Organizations* 13(1):25–48.

Halperin, Morton and Priscilla Clapp. 2006. *Bureaucratic Politics and Foreign Policy*. Washington D.C.: Brookings Institution Press.

Irwin, Douglas A. 2017. *Clashing Over Commerce: A History of US Trade Policy*. Chicago: University of Chicago Press.

Katzenstein, Peter J. 1985. *Small States in World Markets: Industrial Policy in Europe*. Ithaca: Cornell University Press.

Kuk, John, Deborah Seligsohn and Jiakun Jack Zhang. 2022. "The partisan divide in U.S. congressional communications after the China shock." *Economics & Politics* 34(3):494–526.

Linardi, Sera and Nita Rudra. 2020. "Globalization and Willingness to Support the Poor in Developing Countries: An Experiment in India." *Comparative Political Studies* 53(10-11):1656–1689.

Lindbeck, Assar and Jörgen W. Weibull. 1987. "Balanced-Budget Redistribution as the Outcome of Political Competition." *Public Choice* 52(3):273–297.

Malis, Matt, B. Peter Rosendorff and Alastair Smith. 2022. "A Political Economy of International Organizations.".

Marshall, Monty G and Ted Robert Gurr. 2020. Polity5: Political Regime Characteristics and Transitions, 1800-2018. Technical report Center for Systemic Peace.

Melnick, Justin and Alastair Smith. 2023. "International Negotiations in the Shadow of Elections." *Journal of Conflict Resolution* 67(7-8):1452–1481.

Milner, Helen V. 2021. "Voting for Populism in Europe: Globalization, Technological Change, and the Extreme Right." *Comparative Political Studies* 54(13):2286–2320.

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.

Pevehouse, Jon CW, Timothy Nordstrom, Roseanne W McManus and Anne Spencer Jamison. 2020. "Tracking organizations in the world: The Correlates of War IGO Version 3.0 datasets." *Journal of Peace Research* 57(3):492–503.

Putnam, Robert D. 1988. "Diplomacy and Domestic Politics: The Logic of Two-Level Games." *International Organization* 42(3):427–460.

Rodrik, Dani. 1997. *Has Globalization Gone Too Far?* Washington, D.C: Peterson Institute for International Economics.

Rodrik, Dani. 2018. "Populism and the economics of globalization." *Journal of International Business Policy* 1(1):12–33.

Rudra, Nita, Irfan Nooruddin and Niccolò W. Bonifai. 2021. "Globalization Backlash in Developing Countries: Broadening the Research Agenda." *Comparative Political Studies* 54(13):2416–2441.

Ruggie, John Gerard. 1982. "International regimes, transactions, and change: embedded liberalism in the postwar economic order." *International Organization* 36(2):379–415.

Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge: Harvard University Press.

Schonfeld, Bryan. 2021. "Trading Places, Trading Platforms: The Geography of Trade Policy Realignment." *International Organization* 75(4):959–990.

Schonfeld, Bryan and Sam Winter-Levy. 2021. "Policy or Partisanship in the United Kingdom? Quasi-Experimental Evidence from Brexit." *The Journal of Politics* 83(4):1450–1461.

von Borzyskowski, Inken and Felicity Vabulas. 2019. "Hello, goodbye: When do states withdraw from international organizations?" *The Review of International Organizations* 14(2):335–366.

Walter, Stefanie. 2021*a*. "The Backlash Against Globalization." *Annual Review of Political Science* 24(1):421–442.

Walter, Stefanie. 2021*b*. "Brexit Domino? The Political Contagion Effects of Voter-endorsed Withdrawals from International Institutions." *Comparative Political Studies* 54(13):2382–2415.

# Appendix

## Formalization of the Game

Leaders $L$ and $H$ choose to remain in or exit from an international agreement, $\rho_d \in \{0,1\}$ where $\rho_d = 1$ signifies remain for $d = L, H$. Leaders also choose redistributive policies $\theta_{da} \in [0,1]$ for $a = \{remain, exit\}$. Share $\theta_{da}$ represents the amount of the economy allocated to winners: leaders decide whether to stay or leave and how much they want to redistribute. A strategy for leader $d$ is a choice $\rho_d \in \{0,1\}$ and proposals $\theta_{da} \in [0,1]^2$.

Voters observe leaders' proposals and decide whether to vote for $L$ or $H$. Voters are a continuum $i \in [0,1]$ and are either winners or losers. There are $1-\lambda$ winners and $\lambda$ losers, $\lambda \in [0,1]$. Voters receive per capita payoffs based on the income allocated to their group by leader $d$. Voters have increasing and concave payoffs $v(\cdot)$ over income. For simplicity, use logarithmic utility, $v(x) = \log(x)$. Then, for example, if $L$ were to win the election having proposed to remain in the agreement, the income payoff to winners would be $\log(\frac{\gamma\theta_{LR}}{1-\lambda})$ and would be $\log(\frac{\gamma(1-\theta_{LR})}{\lambda})$ to losers. Payoffs are more thoroughly defined in Table 1 of the text.

Voter $i$ in group $j$ also receives a common shock $\beta \sim U[-\frac{1}{2b}, \frac{1}{2b}]$ and an individual shock $\mu_i \sim U[-\frac{1}{2m}, \frac{1}{2m}]$ "in favor" of $L$. Therefore, given $L$'s proposal $\theta_{La}$ and $H$'s proposal $\theta_{Ha}$, voter $i$ votes for $L$ if and only if $v_{ij}(\theta_{La}) + \beta + \mu_i \geq v_{ij}(\theta_{Ha})$. A voter's strategy is a choice to vote for $L$ or $H$ given leaders' choices to remain or exit and accompanying proposals $\theta_{da}$, as well as the realized valence shocks. Formally, $\sigma_{ij}$ : $\{0,1\}^2 \times [0,1]^4 \times [-\frac{1}{2m}, \frac{1}{2m}] \times [-\frac{1}{2b}, \frac{1}{2b}] \to \{L, H\}$.

Leaders choose $\theta_{da}$ to maximize their electoral success. Define $D_j = v_{ij}(\theta_{La}) - v_{ij}(\theta_{Ha})$. We construct $\pi(D_w, D_\ell)$ as the probability that $L$ wins the election as a function of proposals $\theta_{da}$ in four different scenarios: 1. $L$ and $H$ both remain in the agreement, 2. $L$ stays, $H$ exits, 3. $L$ exits, $H$ stays, 4. $L$ and $H$ both exit. With some abuse of notation, index $\pi(\cdot,\cdot)$ by these four scenarios. Then we can express expected utilities as

$$
\begin{aligned}
EU_L(\rho_L, \theta_{LR}, \theta_{LE}; \rho_H, \theta_{HR}, \theta_{HE}) = \; & \rho_L \Big[ \underbrace{\rho_H \Psi \pi_1}_{L \text{ in, } H \text{ in}} + \underbrace{(1-\rho_H)\Psi\pi_2}_{L \text{ in, } H \text{ out}} - \frac{1}{2}(\alpha_R - \theta_{LR})^2 \Big] \\
& + (1-\rho_L)\Big[ \underbrace{\rho_H \Psi \pi_3}_{L \text{ out, } H \text{ in}} + \underbrace{(1-\rho_H)\Psi\pi_4}_{L \text{ out, } H \text{ out}} - \frac{1}{2}(\alpha_E - \theta_{LE})^2 \Big].
\end{aligned}
$$

$$EU_H(\rho_H, \theta_{HR}, \theta_{HE}; \rho_L, \theta_{LR}, \theta_{LE}) = \rho_H \Big[ \underbrace{\rho_L \Psi(1-\pi_1)}_{L \text{ in}, H \text{ in}} + \underbrace{(1-\rho_L)\Psi(1-\pi_3)}_{L \text{ out}, H \text{ in}} - \frac{\kappa}{2}(\alpha_R - \theta_{HR})^2 \Big]$$
$$+ (1-\rho_H) \Big[ \underbrace{\rho_L \Psi(1-\pi_2)}_{L \text{ in}, H \text{ out}} + \underbrace{(1-\rho_L)\Psi(1-\pi_4)}_{L \text{ out}, H \text{ out}} - \frac{\kappa}{2}(\alpha_E - \theta_{HE})^2 \Big].$$

The solution concept is subgame perfect equilibrium. At each of four information sets, voters determine when they would vote for $L$ versus $H$. Given these retention rules, leaders optimally select redistributive proposals $\theta_{da}$ and whether to remain in or exit the agreement.

## Proofs

**Lemma 1** *The probability that $L$ wins the election can be expressed as*

$$\pi\Big(D_w(\theta_{La}, \theta_{Ha}), D_\ell(\theta_{La}, \theta_{Ha})\Big) = \frac{1}{2} + b\Big((1-\lambda)D_w + \lambda D_\ell\Big).$$

**Proof of Lemma 1:** Voter $i$ in group $j$ votes for $L$ whenever $D_j + \mu_i + \beta \geq 0$, where $D_j$ is the difference in income from what $L$ proposes versus what $H$ proposes. Alternatively, voter $i$ votes for $L$ if $\mu_i \geq -D_j - \beta$. Then, $P(\mu_i \geq -D_j - \beta) = \frac{1}{2} + m(D_j + \beta)$. Hence the fraction of winners supporting $L$ is $(1-\lambda)\Big(\frac{1}{2} + m(D_w + \beta)\Big)$ and the fraction of losers supporting $L$ is $\lambda\Big(\frac{1}{2} + m(D_\ell + \beta)\Big)$. To win the election, $L$ must have support satisfying $(1-\lambda)\Big(\frac{1}{2} + m(D_w + \beta)\Big) + \lambda\Big(\frac{1}{2} + m(D_\ell + \beta)\Big) \geq \frac{1}{2}$. This occurs when $\beta \geq -\Big((1-\lambda)D_w + \lambda D_\ell\Big)$. Finally, $P\Big(\beta \geq -\Big((1-\lambda)D_w + \lambda D_\ell\Big)\Big) = \frac{1}{2} + b\Big((1-\lambda)D_w + \lambda D_\ell\Big)$ as in the lemma. ■

**Proposition 1** *The equilibrium proposal by leader $d$ in outcome $a$ solves*

$$b\Psi \frac{\lambda}{1-\theta^*_{da}} = b\Psi \frac{1-\lambda}{\theta^*_{da}} + (\alpha_a - \theta^*_{da})\kappa_d.$$

**Proof of Proposition 1:** Each share $\theta_{da}$ maximizes leader $d$'s utility of taking action $a$ as a best response to the other leader's behavior. All four choice variables – $\theta_{LR}$, $\theta_{LE}$, $\theta_{HR}$, and $\theta_{HE}$ – solve the same type of problem, so it suffices to derive the first-order condition for one choice and generalize accordingly. Let $p_d$ be the probability that leader $d$ exits. Consider $H$'s proposal of national income when remaining in the agreement, $\theta_{HR}$. This solves

$$\theta^*_{HR} = \text{argmax}_{\theta_{HR}}\ (1-p_L)(1-\pi_1) + p_L(1-\pi_3) - \frac{\kappa}{2}(\alpha_R - \theta_{HR})^2,$$

where $\pi_1$ and $\pi_3$ are determined by Lemma 1 given the proposals $(\theta_{LR}, \theta_{HR})$ and $(\theta_{LE}, \theta_{HR})$ respectively. Differentiating with respect to $\theta_{HR}$ yields

$$(\alpha_R - \theta_{HR})\kappa - b\Psi\Big(\frac{\lambda}{1-\theta_{HR}} - \frac{1-\lambda}{\theta_{HR}}\Big) = 0.$$

Generalizing notation yields the equation in the proposition. The point that satisfies this equation at equality, $\theta^*_{HR}$, is guaranteed to be a maximum, as leaders' utility functions are globally concave. The second-order condition confirms this:

$$-\kappa - b\Psi\Big(\frac{1-\lambda}{\theta^2_{HR}} + \frac{\lambda}{(1-\theta_{HR})^2}\Big) < 0.$$

■

**Corollary 1** *L redistributes more than H,* $\theta^*_{La} \leq \theta^*_{Ha}$.

**Proof of Corollary 1:** Since the cross-partial $\frac{\partial^2 u}{\partial \theta_{HR} \partial \kappa} = \alpha_R - \theta_{HR} \geq 0$, by the implicit function theorem $\frac{\partial \theta^*_{HR}}{\partial \kappa} \geq 0$. Because $\kappa_H = \kappa > 1 = \kappa_L$, we have that $\theta^*_{La} < \theta^*_{Ha}$. ■

**Corollary 2** *Equilibrium proposals* $\theta^*_{dR}$ *are increasing in the winners' gains* $\alpha_R$.

**Proof of Corollary 2:** Since the cross-partial $\frac{\partial^2 u}{\partial \theta_{HR} \partial \alpha_R} = \kappa > 0$, by the implicit function theorem $\frac{\partial \theta^*_{HR}}{\partial \alpha_R} > 0$. ■

**Proposition 2** *There exist thresholds* $\gamma_L \leq \gamma_H$ *such that leader d exits the agreement whenever* $\gamma \leq \gamma_d$ *and remains otherwise.*

**Proof of Proposition 2:** We will establish existence of $\gamma_H$, the proof for $\gamma_L$ is analogous. Let $p_d$ be the probability that leader $d$ exits. $H$'s indirect utility functions for remaining and exiting are

$$\begin{aligned}
EU_H(remain) &= \frac{1}{2}(\Psi - \kappa(\alpha_R - \theta^*_{HR})^2) + b\Psi\Big(p_L\lambda\log(\frac{\theta^*_{LE}(1-\theta^*_{LR})}{\theta^*_{LR}(1-\theta^*_{LE})}) + p_L\log(\frac{\gamma\theta^*_{LR}}{\theta^*_{LE}}) \\
&+ \log(\frac{\theta^*_{HR}}{\theta^*_{LR}}) + \lambda\log(\frac{\theta^*_{LR}(1-\theta^*_{HR})}{\theta^*_{HR}(1-\theta^*_{LR})})\Big). \\
EU_H(exit) &= \frac{1}{2}(\Psi - \kappa(\alpha_E - \theta^*_{HE})^2) + b\Psi\Big(p_L\lambda\log(\frac{\theta^*_{LE}(1-\theta^*_{LR})}{\theta^*_{LR}(1-\theta^*_{LE})}) + p_L\log(\frac{\gamma\theta^*_{LR}}{\theta^*_{LE}}) \\
&+ \log(\frac{\theta^*_{HE}}{\gamma\theta^*_{LR}}) + \lambda\log(\frac{\theta^*_{LR}(1-\theta^*_{HE})}{(1-\theta^*_{LR})\theta^*_{HE}})\Big).
\end{aligned}$$

Therefore, $H$ prefers to remain whenever

$$\Phi_H(\gamma) := \frac{\kappa}{2}(\alpha_E + \alpha_R - \theta^*_{HE} - \theta^*_{HR})(\alpha_E - \alpha_R - \theta^*_{HE} + \theta^*_{HR}) + b\Psi\Big(\log(\frac{\gamma\theta^*_{HR}}{\theta^*_{HE}}) + \lambda\log(\frac{\theta^*_{HE}(1-\theta^*_{HR})}{\theta^*_{HR}(1-\theta^*_{HE})})\Big) > 0.$$

Since $\Phi_d(\gamma)$ is strictly increasing in $\gamma$, $\frac{\partial\Phi_d}{\partial\gamma} = \frac{b\Psi}{\gamma} > 0$, by the intermediate value theorem there is a point $\gamma_d$ where $EU_d(remain) = EU_d(exit)$ with $EU_d(exit) > EU_d(remain)$ whenever $\gamma < \gamma_d$.

To rank, we demonstrate that $\frac{\partial\gamma_H}{\partial\kappa} \geq 0$. Differentiating yields $\frac{\partial\Phi_H}{\partial\kappa} = \frac{1}{2}\Big(\alpha_E^2 - \alpha_R^2 - 2\alpha_E\theta^*_{HE} + \theta^{*^2}_{HE} + 2\alpha_R\theta^*_{HR} - \theta^{*^2}_{HR}\Big) \leq 0$. Therefore by the implicit function theorem, $\frac{\partial\gamma_H}{\partial\kappa} \geq 0$. Since $\kappa_H = \kappa > 1 = \kappa_L$, $\gamma_H > \gamma_L$. ■

**Corollary 3** *The thresholds $\gamma_d$ are increasing in the gains to winners $\alpha_R$.*

**Proof of Corollary 3:** We compute $\frac{\partial\gamma_d}{\partial\alpha_R} = -\frac{\partial\Phi_d/\partial\alpha_R}{\partial\Phi_d/\partial\gamma}$. From Proposition 2, $\frac{\partial\Phi_d}{\partial\gamma} > 0$. By the envelope theorem, the first term of $\frac{d\Phi_d}{d\alpha_R} = \frac{\partial\Phi_d}{\partial\theta_{da}}\frac{\partial\theta_{da}}{\partial\alpha_R} + \frac{\partial\Phi_d}{\partial\alpha_R}$ is zero. Differentiating, $\frac{\partial\Phi_d}{\partial\alpha_R} = (\theta_{dR} - \alpha_R)\kappa_d < 0$. Then, by the implicit function theorem, $\frac{\partial\gamma_d}{\partial\alpha_R} \geq 0$. ■

**Proposition 3** *Increasing the winners' gains from globalization $\alpha_R$ can create political realignments.*

**Proof of Proposition 3:** It is sufficient to demonstrate that there exist cases in which increasing $\alpha_R$ moves the equilibrium outcome from both $L$ and $H$ remaining to $L$ remaining and $H$ exiting, and that $D^*_w$ and $D^*_\ell$ change sign. Consider a case where $\lambda = 0.66$, $\alpha_E = 0.4$, $b = 1$, $\Psi = 1$, and $\kappa = 2.5$. Let $\alpha_R = 0.64$. Computed equilibrium shares are $\theta^*_{LR} = 0.439$, $\theta^*_{LE} = 0.359$, $\theta^*_{HE} = 0.372$, and $\theta^*_{HR} = 0.507$. We then calculate $\gamma_H = 1.078$ and $\gamma_L = 1.04$. Pick $\gamma = 1.08$ so that neither $H$ nor $L$ exit. Then $D^*_w = -0.143$, so winners support $H$, and $D^*_\ell = 0.129$, so losers support $L$.

Now increase $\alpha_R = 0.65$. Equilibrium proposals are $\theta^*_{LR} = 0.442$, $\theta^*_{LE} = 0.359$, $\theta^*_{HE} = 0.372$, and $\theta^*_{HR} = 0.512$. With these shares, $\gamma_H = 1.084$ and $\gamma_L = 1.043$, so $H$ prefers to exit but $L$ does not. Then $D^*_w = 0.249$, so winners support $L$, and $D^*_\ell = -0.041$, so losers support $H$, establishing existence. ■