

Why Exit from International Agreements? A Domestic Perspective*

Justin Melnick[†]

May 6, 2023

Abstract

What explains withdrawal from international agreements? I argue that the implementation of redistributive policies intended to compensate globalization “losers” varies by the costs that leaders face to making policy. Leaders with high policymaking costs may have political incentives to propose exit from agreements rather than maintain the globalized status quo with insufficient levels of redistribution. Exit from agreements emerges in equilibrium because these candidates cannot credibly promise to offset losers’ incurred damages with redistributive policy. The model demonstrates that rising inequality between globalization winners and losers increases the likelihood of exit and subsequently makes redistribution more politically costly. Rising inequality generates realignments in political support such that losers may *ex-ante* demand compensation to sustain globalization, but winners support candidates less likely to supply it; if exit is credible, winners *ex-post* would have preferred to compensate but losers no longer view redistribution as sufficient. The theory thus explicates political incentives to withdraw from agreements despite globalization being welfare-maximizing.

Word Count: 7846 (text) + 1396 (appendix) = 9234 (total)

*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, and GSIPE for insightful comments and feedback.

[†]Ph.D. Student, Department of Politics, New York University. Contact: melnickj@nyu.edu and [justinmelnick.github.io](https://github.com/justinmelnick).

Redistribution from globalization’s “winners” to its “losers” is central to the social compact that sustains international integration (Cameron 1978; Ruggie 1982; Katzenstein 1985; Rodrik 1998). While some adjustment has occurred (Kim and Pelc 2021*b*), redistributive policies have in general failed to sufficiently compensate the losers (Hays 2009; Bonica et al. 2013; Egger, Nigai and Strecker 2019). Concomitant rising inequality requires greater redistribution to offset greater losses, but no politically feasible coalition has developed in which winners are willing to finance larger transfers (Linardi and Rudra 2020) that losers are willing to accept (Bowen, Broz and Rosendorff 2022). Recent empirical work finds that globalization’s losers no longer find compensation sufficient, preferring anti-integration policies instead (Flaherty and Rogowski 2021; Milner 2021). Consequently, electoral challengers have emerged hoping to assuage losers through means other than domestic compensation, instead promoting modifications to or withdrawal from international institutions (Bornschier 2017; Colantone and Stanig 2018; 2019).

The dearth of domestic redistribution has contributed to a political “globalization backlash” (Mansfield, Milner and Rudra 2021; Walter 2021*a*). This backlash is particularly puzzling given that globalization is welfare-enhancing. Why have governments failed to allocate globalization’s surplus in a way that makes everyone better off? Why would leaders pursue anti-globalization policies that carry a deadweight loss for their electorates – and why would voters support those leaders? In this paper, I provide a formal model to explain the empirical regularities of the backlash and to illustrate how leaders can find it politically optimal to withdraw from international agreements despite globalization being welfare-enhancing. I propose a novel mechanism to resolve this puzzle, focusing on the differential costs leaders face in implementing redistributive policies.

Leaders vary in their costs of “getting things done,” or implementing policies that differ from the status quo. Differences in these costs drive politicians’ preferences for redistribution or exit. High policymaking costs generate redistributive frictions. Therefore, leaders who incur high costs to changing redistributive policy may have political incentives to propose exit from agreements rather than maintain the globalized status quo with insufficient levels of redistribution. Moreover, the introduction of anti-globalization candidates induces a realignment in political coalitions of whom winners and losers are likely to support electorally. When some leaders prefer to exit but others cannot credibly to do, the “traditional” candidate of the winners forsakes globalization in order to win the support of the losers.

Main Results

I propose a formal model that studies a domestic political contest over globalization policy. Globalization produces “winners” and “losers” domestically. Due to sectoral specialization, winners command greater per capita shares of national income than losers. Leaders can propose a transfer of income from winners to losers to fund compensation programs that maintain the globalized status quo, or they can choose to withdraw from the international agreement. Withdrawal is inefficient and contracts the size of the pie.

Two candidates compete for political office who vary in their ability to “get things done,” or move policy away from the status quo. The election features candidate L , a politician with “low” policymaking costs, against candidate H , a politician with “high” policymaking costs. We can think of these costs as features intrinsic to candidates, like their competence or ability to effectively manage government. Alternatively, policymaking costs may be consequences of the fact that leaders from different parties appoint bureaucrats with distinct policy preferences, so high policymaking costs would stem from bureaucratic resistance to implementing redistributive policies. Higher costs of policymaking generate discrepancies in the redistributive proposals that leaders advance. These policymaking frictions imply that, when designing transfer schemes, candidates with high policymaking costs are biased toward the status quo income distribution: this bias naturally ingratiates high-cost leaders with the globalization winners. Globalization losers, by contrast, prefer a candidate who can more efficiently redistribute away from the status quo. However, high policymaking costs also generate a willingness to withdraw from international agreements. In so doing, high-costs candidates realign themselves with globalization losers; losers abandon their appetite for compensation in favor of a larger share of a smaller pie. Exit therefore emerges in equilibrium because candidates with high policymaking costs cannot credibly promise to offset losers’ incurred damages with redistributive policy.

Inequality serves as an essential moderator in observing support for exit. Office-seeking politicians will move away from compensation-based policy responses and toward exit from international agreements as globalization-caused inequality increases. I show that increasing the share of globalization losers makes compensation harder to sustain politically: as inequality grows, so too do the costs of redistributing from winners to losers. Leaders need to promise more to the losers to maintain the globalized status quo, thereby taxing winners more heavily. Additionally, rising inequality incentivizes politicians to withdraw rather than continue globalization; voters subsequently select candidates who are willing to exit from international institutions rather than uphold the system of international liberalization.

The model highlights the delicate conditions under which embedded liberalism can be expected to subsist. When exit is not a credible policy option for either candidate, winners prefer leaders with higher policymaking

costs who subsequently redistribute less, while losers demand greater compensation by supporting candidates who can redistribute more efficiently. If withdrawal becomes credible, for example by increasing the share of globalization losers, then losers break with their support of the pro-redistribution candidate in favor of the candidate promising exit. Conversely, winners switch their political allegiances to the candidate with low policymaking costs because, despite redistribution, they would be better off under a system of globalization. Therefore, losers *ex ante* demand the compensation to sustain globalization yet winners support candidates less likely to supply it; once exit emerges as a plausible policy, winners *ex post* would have preferred to compensate but losers no longer view redistribution as sufficient.

Domestic political consequences of globalization affect the propensity for withdrawal from international agreements. Consider the fate of the Trans-Pacific Partnership (TPP) in the 2016 United States presidential election. Hillary Clinton, a known supporter of pro-globalization policies and one of the architects of the agreement, advocated for the TPP as the new “gold standard” in crafting trade deals. Meanwhile, Donald Trump argued that “the TPP would be the death blow for American manufacturing. It would give up all of our economic leverage to an international commission that would put the interests of foreign countries above our own.”¹ Trump criticized Clinton’s support of the agreement and appealed to the anticipated losses of American workers and manufacturers, traditionally seen as globalization losers in the United States, and an integral component of the Democratic coalition. As the Republican nominee, Trump aligned himself with a party with a distaste for redistribution, suggesting that his administration would face greater costs in implementing redistributive policies relative to Clinton’s. The expectation of less redistribution rendered Trump’s embrace of policy alternatives like withdrawal from the TPP more credible. Ultimately, the domestic electoral pressures surrounding the agreement even compelled Clinton to revoke support of the TPP, quashing any prospects for its passage, and further reflecting how domestic political concerns are central to the implementation of international agreements. Trump’s embrace of withdrawal as a policy measure exemplifies that, through processes of domestic political competition, enterprising political candidates can seize on the opportunity to form an electoral coalition of globalization losers, forsaking gains from international cooperation for domestic political expedience.

Contribution

This paper provides two principal contributions. First, I introduce a simple theoretical model to explain how domestic policymaking frictions serve as a commitment device to credibly withdraw from international

¹<https://www.politico.com/story/2016/06/full-transcript-trump-job-plan-speech-224891>

agreements. 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 of globalization policy. High policymaking costs introduce frictions into the redistributive process and subsequently define the scope of the political appetite for redistribution.

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 a small, “elite” faction of society, conditions support for anti-globalist or populist candidates. They document that rising inequality is necessary to elicit voter demand for leaders promising policies like exit. [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.

Scholars have extensively investigated the domestic forces that shape the creation of international agreements (e.g., [Putnam 1988](#); [Milner and Rosendorff 1997](#); [Buisseret and Bernhardt 2018](#); [Melnick and Smith 2022](#)), but we have yet to understand how domestic politics affects leaders’ decision to exit from such agreements. 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 absolve membership in 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.² Increased incidence of exit may also be due to contagion effects (e.g., [Walter 2021b](#)), in which withdrawal by one state motivates others to follow suit, in accordance with 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. Finally, a strand of the literature has considered exit as a consequence of growing regime complexity and bounded rationality, in which exit from some international commitments becomes inevitable when they

²While 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.

are superseded by less constricting forms of integration (Haftel and Thompson 2018; Ge 2022).

To date, this paper provides one of the first microfounded accounts of exit from international agreements. Policy frictions explain variation in candidates’ proposals of pro- and anti-globalization policies. This drives leaders to commit to exit as a means of shifting electoral outcomes, despite being welfare inferior. Put differently, when the political costs of domestic distributional conflict outweigh benefits of international cooperation (Rodrik 2018; Flaherty and Rogowski 2021), leaders abandon globalization in an attempt to preserve their electoral fortunes.

Model Setup

I study the downstream, domestic consequences of membership in an international agreement. Within this country, the agreement fosters “winners” and “losers.” Let the domestic public be a continuum with each individual indexed by $i \in [0, 1]$ and let some share $\lambda \in [0, 1]$ of these individuals be globalization losers (ℓ) while $1 - \lambda$ are globalization winners (w). We can interpret this cleavage such that a fraction $1 - \lambda$ of the Home population is employed in sectors that benefit from the agreement, while a share λ is adversely affected. The value of λ is common knowledge.

Winners and losers command different shares of the economy based on their sectoral specializations. Normalize the size of the economy to 1 under the agreement. In the status quo, winners own a share $\alpha_R \in [0, 1]$ of the economy and losers own $1 - \alpha_R$. The parameter α_R is an exogenous, status quo point for policy under globalization and represents the structural, macroeconomic factors that affect returns to income for winners and losers. What makes the $1 - \lambda$ individuals “winners” is that their per capita gains in the agreement are greater than the losers, or that globalization makes some domestic groups better off than others (Stolper and Samuelson 1941). We therefore restrict attention to the case where $\frac{\alpha_R}{1-\lambda} > \frac{1-\alpha_R}{\lambda}$ or $\lambda > 1 - \alpha_R$, which encodes the fact that in a globalized status quo with zero redistribution, winners have higher per capita income than losers.

I refer to the share of losers λ as domestic “inequality” created by globalization. This is a reduced-form operationalization of inequality that is purely compositional. The primitive λ captures the relative size of the winners versus the losers: increasing λ means that the *ex ante* gains from globalization α_R are concentrated in the hands of fewer citizens. It does not capture differences in the groups’ per capita incomes. Such a measure of societal income inequality would be defined as the ratio of winners’ and losers’ incomes, $\eta(x) = \frac{x}{1-\lambda} / \frac{1-x}{\lambda}$ for some share x of national income owned by winners. In equilibrium, income inequality

η is increasing in λ , so it is without loss of generality to think of increasing λ as increasing inequality across winners and losers. I will therefore use “inequality” and the “share of losers” to refer to λ interchangeably.

The game depicts a domestic political contest between two candidates, L (she) and H (he). Candidates seek to propose globalization policy that will maximize their prospects of winning office. Each candidate’s platform consists of two elements. L and H simultaneously decide whether to remain in or exit from the agreement, as well as how to transfer income between winners and losers under either international policy outcome. I employ a common framework of electoral competition in which candidates make binding campaign policy choices (cf. [Downs 1957](#)) so that we can study the electoral incentives of proposing to remain in or exit from an international agreement.

Exit is inefficient. If the agreement is abrogated, then the economy contracts to size $\gamma < 1$. In addition, the status quo point for policy contracts to $\alpha_E \leq \alpha_R$ and serves as the new income distribution between winners and losers against which candidates propose any transfers. This assumption reflects the fact that, prior to any government intervention, income accruing to winners is greater when the country is party to the agreement.

Leaders propose policies that may redistribute income away from the exogenous macroeconomic shares α_R or α_E . To do so, they consider how the enactment of different policies affects their chances of winning the election, given what the other candidate would propose, as well as the cost of changing policy. Candidates vary in the extent to which they find it costly to change policy away from the status quo, which I refer to as their *costs of policymaking*. We suppose that L experiences “low” costs of policymaking, which may imply that she intrinsically is a competent manager of government or technocrat, or is from a party with preferences for redistributive policies so she can staff an administration with pro-redistribution bureaucrats. L ’s costs of policymaking are normalized to 1. Conversely, H is a politician with “high” costs of policymaking, which may reflect greater difficulty in motivating his party’s bureaucratic appointees to enact redistributive policies. H experiences a cost $\kappa > 1$ to move policy away from the status quo. The parameter κ is common knowledge.

Each candidate chooses a share of national income to be allocated to winners, θ_{da} for $d \in \{L, H\}$ and $a \in \{remain, exit\}$. A strategy for each candidate is to propose whether to remain in or exit from the agreement, and, how to divide the pie in each of these two possible international policy outcomes.

After L and H have announced their proposals, citizens go to the polls. Let voters have concave payoffs $v(\cdot)$ over income, which we specify using a logarithmic utility function, $v(x) = \log(x)$. The following table summarizes the per capita income to winners and losers for each candidate’s proposals under the two possible international policy outcomes.

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

Table 1: Income Distribution across Winners and Losers

Voters compare the differences in their expected income under L versus H , and also take into account an individual-specific shock μ_{ij} as well as a common preference shock β . Let $\mu_{ij} \sim U[-\frac{1}{2m_j}, \frac{1}{2m_j}]$ and let $\beta \sim U[-\frac{1}{2b}, \frac{1}{2b}]$. The parameters m_j and b define the salience of globalization policy relative to other issues in the electoral landscape. Define D_j as the difference in expected income, $D_j = v(L) - v(H)$. For example, if both L and H commit to remaining in the agreement, then $D_w = \log(\frac{\theta_{LR}}{1-\lambda}) - \log(\frac{\theta_{HR}}{1-\lambda})$ and $D_\ell = \log(\frac{1-\theta_{LR}}{\lambda}) - \log(\frac{1-\theta_{HR}}{\lambda})$. Voter i in group $j \in \{w, \ell\}$ will prefer L over H whenever

$$D_j + \mu_{ij} + \beta \geq 0.$$

Leaders care exclusively about their electoral prospects. When proposing policies, L and H maximize their chances of winning the election less the cost of policymaking. If elected, leaders enjoy a benefit $\Psi > 0$ and receive a payoff of zero if they lose the election. If $\pi(D_w, D_\ell)$ is the probability that L wins the election (to be derived), then candidates choose θ_{da} to maximize

$$\begin{aligned} EU_L(\theta_{La}, \theta_{Ha}) &= \pi\left(D_w(\theta_{La}, \theta_{Ha}), D_\ell(\theta_{La}, \theta_{Ha})\right)\Psi - \frac{1}{2}(\alpha_a - \theta_{La})^2. \\ EU_H(\theta_{La}, \theta_{Ha}) &= \left(1 - \pi\left(D_w(\theta_{La}, \theta_{Ha}), D_\ell(\theta_{La}, \theta_{Ha})\right)\right)\Psi - \frac{\kappa}{2}(\alpha_a - \theta_{Ha})^2. \end{aligned}$$

Based on the optimal θ_{da} s, candidates subsequently determine whether to remain in or exit from the agreement.

To recapitulate, the sequence of the game is as follows.

1. Candidates L and H simultaneously announce intentions to remain or exit the agreement, proposing divisions of national income θ_{da} .
2. Valence shocks μ_{ij} and β are realized. An election occurs.
3. The election winner's policy outcome is implemented. Payoffs are realized. Game ends.

We solve for the subgame perfect equilibrium of the game via backward induction. A more formal definition of the game is in the appendix.

Analysis

We first derive the probability that L wins the election $\pi(D_w, D_\ell)$ based on voter behavior, and then consider how L and H divide national income and determine optimal globalization policy. We then state the equilibrium, in which candidates prefer to exit the agreement only when it is not too inefficient. All proofs are in the appendix.

There are four scenarios that voters face. Both candidates could propose to remain, 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 candidate, as in Table 1. A voter i in group j votes for L whenever $v(L) + \mu_{ij} + \beta \geq v(H)$, or when $\underbrace{v(L) - v(H)}_{D_j} + \mu_{ij} + \beta \geq 0$. Then, for any values of D_w or D_ℓ , we obtain a general expression for the probability that L wins the election, stated in the following lemma.³

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

$$\pi(D_w, D_\ell) = \frac{1}{2} + b\left((1 - \lambda)D_w + \lambda D_\ell\right).$$

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. If voters' decision rules were completely insensitive to globalization policy, formally $b \rightarrow 0$, then the election would be decided on valence shocks alone, which in expectation have value zero. Hence, L would win the election with probability $\frac{1}{2}$.

Equilibrium Redistribution

L and H simultaneously announce their intentions to remain in or exit from the agreement, as well as how they would divide national income among winners and losers in each case. Our first result concerns the characterization and rank ordering of equilibrium transfer proposals θ_{da}^* . Candidates balance the costs of

³Without loss of generality, we will assume $m_w = m_\ell = m$. The assumption states that the support of the valence shocks is the same for winners and losers, or that the groups view globalization as equally salient. A more general expression is in the appendix.

polymaking associated with redistribution and the marginal benefit that redistribution has in advancing a candidate's electoral success.

Proposition 1 *The equilibrium division of national income by candidate d in proposing outcome a solves*

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

Proposals can be ranked such that $\theta_{LE}^ \leq \theta_{HE}^* \leq \theta_{LR}^* \leq \theta_{HR}^*$.*

Polymaking costs induce redistributive frictions. Since it is more costly for H to change policy, his proposals are biased toward the status quo. By contrast, low-cost L can redistribute income more efficiently, providing more income to the losers. It follows that $\theta_{La}^* \leq \theta_{Ha}^*$, or that L always proposes greater redistribution than H . Under both policy outcomes, both candidates redistribute income from winners to losers, but H retains a greater share of income for the winners than L . This clearly generates an induced political preference among winners for H , while losers prefer L .

Candidates' proposals are also sensitive to compositional shifts. Quite intuitively, increasing the share of losers requires candidates to redistribute more. As such, compensation becomes more costly, as it detracts income away from the winners.

Corollary 1 *Equilibrium proposals θ_{da}^* are:*

- *decreasing in the share of losers λ ;*
- *increasing in the status quo share to winners α_a ;*
- *increasing in polymaking costs κ .*

Corollary 1 also formalizes the idea that politicians with high polymaking costs fail to adequately redistribute and are biased toward providing greater shares of income to the beneficiaries of the status quo, the globalization winners. When the costs of polymaking are large, leaders can do little to shift away from the status quo. Finally, transfers are increasing in the status quo point, which simply means that if winners start out with a greater share of the pie, their post-transfer share is increasing as well.

Note that despite the fact that leaders optimally equalize the marginal cost of polymaking with the marginal benefit of redistribution to choose θ_{da}^* , the returns to redistribution themselves have electoral costs by pitting winners against losers. Transfers attempt to sway enough losers into remaining in a candidate's

political coalition, but cannot be too high so as not to alienate the winners. By increasing the share of losers in society, redistribution becomes more expensive since losers demand a greater share of the pie. Hence any compensation away from the point α_a must be paid for through declining political support among winners.

In addition to its political costs, redistribution fails to abate income inequality when the share of losers grows. For any equilibrium redistributive proposal, the ratio of per capita payoffs can be written as $\eta(\theta_{da}^*) = \frac{\lambda \theta_{da}^*}{(1-\lambda)(1-\theta_{da}^*)}$. Intuitively, the direct effect of increasing λ on η is positive: holding fixed any θ_{da}^* allocated to winners, more losers consuming the same $1 - \theta_{da}^*$ slice of the pie means societal income inequality will go up. Put differently, a greater share of losers implies a smaller share of winners, meaning winners' per capita income becomes concentrated in fewer hands. However, leaders redistribute more when there are more losers in society, $\frac{\partial \theta_{da}^*}{\partial \lambda} \leq 0$. Unfortunately, the strategic effect of increasing λ does not counteract the direct effect. Redistribution does not occur “fast enough” to offset changes in the group size across winners and losers.

Corollary 2 *Income inequality $\eta(\theta_{da}^*)$ is increasing in the share of losers λ .*

Increasing compositional inequality λ adversely affects income inequality in the aggregate. This confirms the notion that interpreting λ as a reduced-form measure of inequality comports with expected changes in other measures of inequality.

When does Exit Occur?

Given the optimal divisions of national income, candidates determine whether it is politically profitable to remain in the agreement or to exit. The following proposition establishes that leaders remain when exit is relatively inefficient.

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

Withdrawal shrinks the size of the pie. If exiting from the agreement contracts the economy too much, then the gains from integration supersede any internal distributional conflicts; exit is therefore not a credible policy option. However, if withdrawal is not too damaging to the size of the pie, then office-seeking candidates may find it politically opportunistic to forsake the gains from international cooperation, despite exit being welfare inferior. Indeed, if γ , the size of the pie post-withdrawal, is large enough, then leaders have political incentives to de-globalize.

Importantly, higher policymaking costs imply both fewer transfers from winners to losers as well as an increased chance of exit. The cutpoint γ_H is less than γ_L , meaning that H is more likely to propose

withdrawal than L . High-cost H is less able to make the requisite transfers under a globalized regime, and would rather abrogate the agreement. This benefits globalization losers, who may now have incentives to switch their political loyalties from L to H . By exiting the agreement, the income distribution falls to $\alpha_E \leq \alpha_R$, which shifts the distribution of income toward losers relative to the distribution of income while in the agreement. H 's high policymaking costs tie his hands in implementing redistributive policies, generating credibility to pursue inefficient outcomes like exit, which may be electorally desirable if it sways the political loyalty of globalization losers.

Inequality further threatens the stability of international liberalization. As the share of globalization losers increases, so too do the conditions under which domestic political candidates find it optimal to withdraw from the agreement. This comports with the empirical regularity that winners are increasingly unwilling to share their gains with losers (Linardi and Rudra 2020), or that no politically feasible transfer exists (Bowen, Broz and Rosendorff 2022).

Corollary 3 *The thresholds γ_d are decreasing in the share of losers λ .*

The incentives to exit increase for both candidates when inequality rises. From Corollary 1, redistribution is more costly when there are more globalization losers. By the assumption that $\alpha_R \geq \alpha_E$, the *ex ante* distribution of income under exit is more favorable to losers than under remain. This means that, prior to transfers, losers can expect to enjoy a greater share of income if the agreement were abrogated, and Proposition 1 confirms that candidates propose smaller shares to winners if they exit compared to if they remain. Therefore, increasing λ renders losers more politically salient, incentivizing candidates to allocate a greater share of the economy to them, which can be more easily achieved by withdrawing from the agreement. International cooperation thus falls victim to electoral expedience. The connection between inequality and the proposal of exit is illustrated in Figure 1, which plots the thresholds γ_L (solid line) and γ_H (dotted line) as a function of the share of globalization losers, both of which are decreasing in λ . To the right of each line represents regions of the parameter space where candidates would be willing to exit the agreement.

Propositions 1 and 2 characterize the game's equilibrium. Three possible outcomes obtain, each carrying differing political implications. When candidates choose the same international action, to stay or to remain, redistributive frictions create natural political alliances. If $\gamma < \gamma_H$, both candidates would prefer to remain in the agreement and globalization continues. International cooperation faces no threat, and leaders propose θ_{LR}^* and θ_{HR}^* as transfers to deal with any globalization-related inequality. H 's higher policymaking costs bind him to policy proposals closer to the status quo α_R , so in expectation the winners break for H and

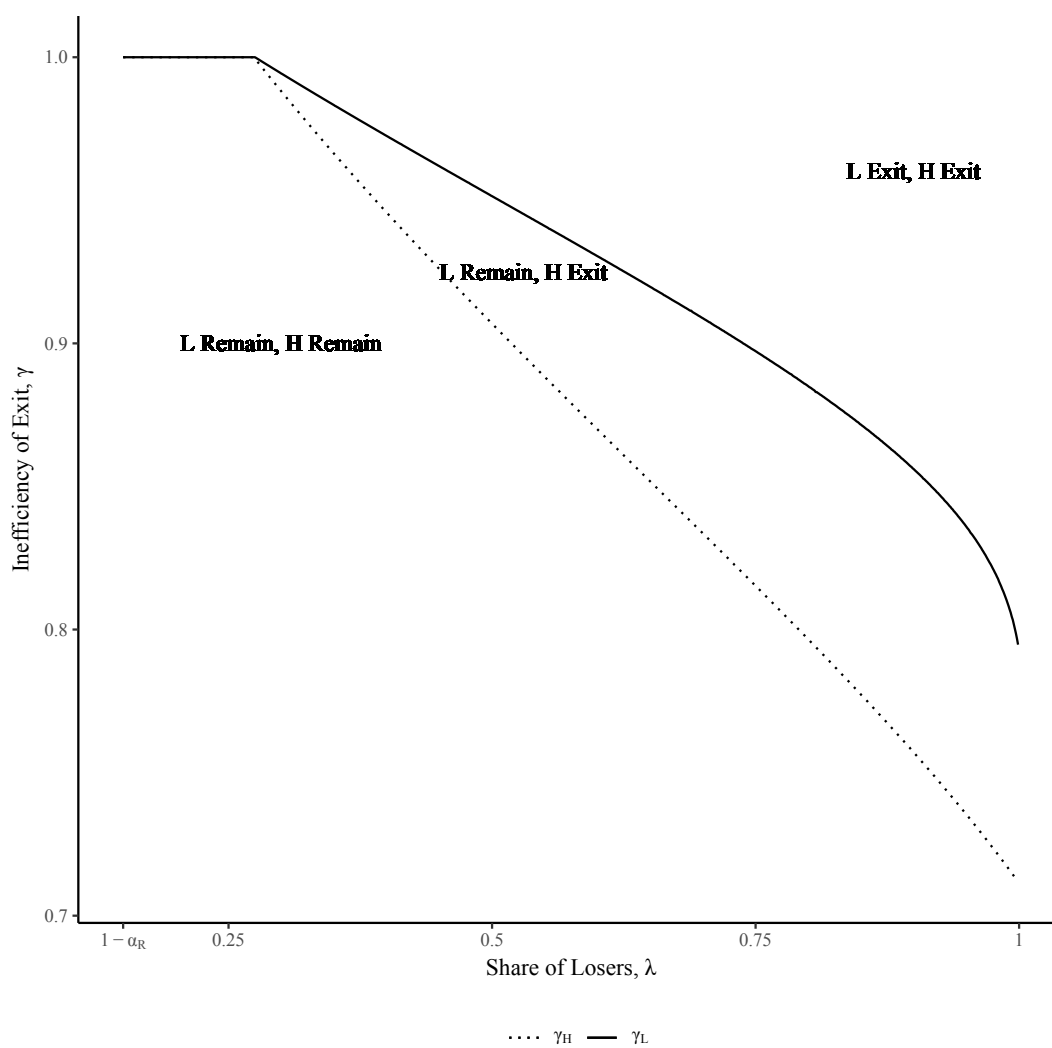


Figure 1: Inequality and Thresholds for Exit by Candidate

the losers rally behind L . Similar behavior occurs if $\gamma > \gamma_L$, when both L and H would exit the agreement, disbursing θ_{LE}^* and θ_{HE}^* . In this region of the parameter space, exit's disturbance costs on the economy are minimal enough that both politicians would abandon globalization. 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.

Electoral coalitions shift when $\gamma_H \leq \gamma \leq \gamma_L$. In this substantively interesting case, globalization remains intact only if L wins the election because H would exit. When voting, losers must decide whether to accept a smaller share of a larger pie by supporting L , or a share close to α_E of a smaller pie by supporting H . Winners consider that H can provide a larger share than L albeit of a smaller pie.

Policy frictions lead to a status quo bias, which implies both fewer transfers and greater willingness to exit on the part of candidates with high policymaking costs. The failure to adequately provide compensation to globalization losers (Walter 2010; Autor, Dorn and Hanson 2013) can thus be explained as the outcome of office-seeking politicians hedging their electoral fortunes on disintegration rather than redistribution. Factors like rising inequality further pressure politicians to implement alternative policies to redistribution, which becomes costlier.

Exit is a rare event (von Borzyskowski and Vabulas 2019). Political support for exit may be observed for equilibrium, but its implementation need not occur if an anti-globalization candidate loses the election. Rather, exit enters the national electoral discourse if it is a politically profitable and credible policy proposal, generating a “backlash” to globalization (Walter 2021a).

The Effects of Globalization-Caused Inequality

We now further explore the implications of rising globalization-caused inequality. As mentioned, inequality exacerbates the tradeoff between remaining in and exiting from the agreement. In particular, we study how rising inequality affects the globalization policies that politicians propose and how changing these policies alters the composition of each candidate's political coalition. Additionally, we consider how candidates' electoral fortunes are affected by a growing share of globalization losers.

Political Realignments

Recall from Proposition 2 above that the game's equilibrium is defined by cutpoints γ_H and γ_L , which delineate candidate willingness to remain in or exit from the agreement. These cutpoints rely on primitives –

such as the share of losers λ – both directly and indirectly through equilibrium redistribution. Corollaries 1 and 3 establish that policy proposals are sensitive to the share of losers in society: both redistribution and the likelihood of exit increase with a growing population of globalization losers. In what follows, we document how increasing λ shifts the equilibrium cutpoints γ_H and γ_L , and therefore changes the regions of the parameter space in which politicians would be willing to exit.

By shifting these cutpoints, rising inequality alters both candidate responses to globalization as well as voters' support for these candidates. Consider a status quo in which neither candidate would be willing to exit from the agreement, defined by $\gamma < \gamma_H$. As the share of globalization losers increases, γ_H decreases: exit becomes a more attractive option for politicians since more potential voters are harmed by globalization. Holding γ constant, it can now be the case that $\gamma_H \leq \gamma \leq \gamma_L$, so that H 's optimal policy proposal now involves exiting from the agreement rather than remaining.⁴

What does this imply for voter behavior? In the status quo, winners benefit from H 's policies because he redistributes less than L , $\theta_{HR}^* \geq \theta_{LR}^*$. However, with rising inequality, and subsequent changes in globalization policy, H 's ability to exit under conditions where L cannot delivers him an entirely new group of political supporters, the globalization losers. Rather than supporting the pro-redistribution L , losers shift their political allegiance to H . For losers, exit delivers a larger share of a smaller pie, formally $\gamma(1 - \theta_{HE}^*) \geq 1 - \theta_{LR}^*$. Conversely, when H switches his optimal policy from remaining to exiting, winners no longer have a pro-globalization yet minimally redistributive candidate. Their next best alternative is L who, despite proposing greater redistribution, allocates winners a larger share of a larger pie, formally $\theta_{LR}^* \geq \gamma\theta_{HE}^*$.

Proposition 3 *Increasing the share of losers can create political realignments.*

The proposition establishes existence of these realignments as differences in *levels* of political support D_j^* as a result of rising inequality, not differences in *changes*. Formally, 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 . Changing the share of losers is required only to move the conditions under which exit becomes optimal, holding γ constant. Put simply, in an equilibrium where both candidates support globalization, losers in expectation support L and winners support H , $D_\ell^* > 0$ and $D_w^* < 0$. However, when we “move” into an equilibrium where H exits but L does not – which can be achieved by increasing the share of losers and subsequently changing the cutpoints that define equilibrium choices to remain or to exit – we now have that $D_\ell^* < 0$ and $D_w^* > 0$, establishing shifts in political support.

⁴Increasing λ further will lead to a decrease in the cutpoints such that $\gamma > \gamma_L$, so both candidates would exit in an equilibrium with these primitives.

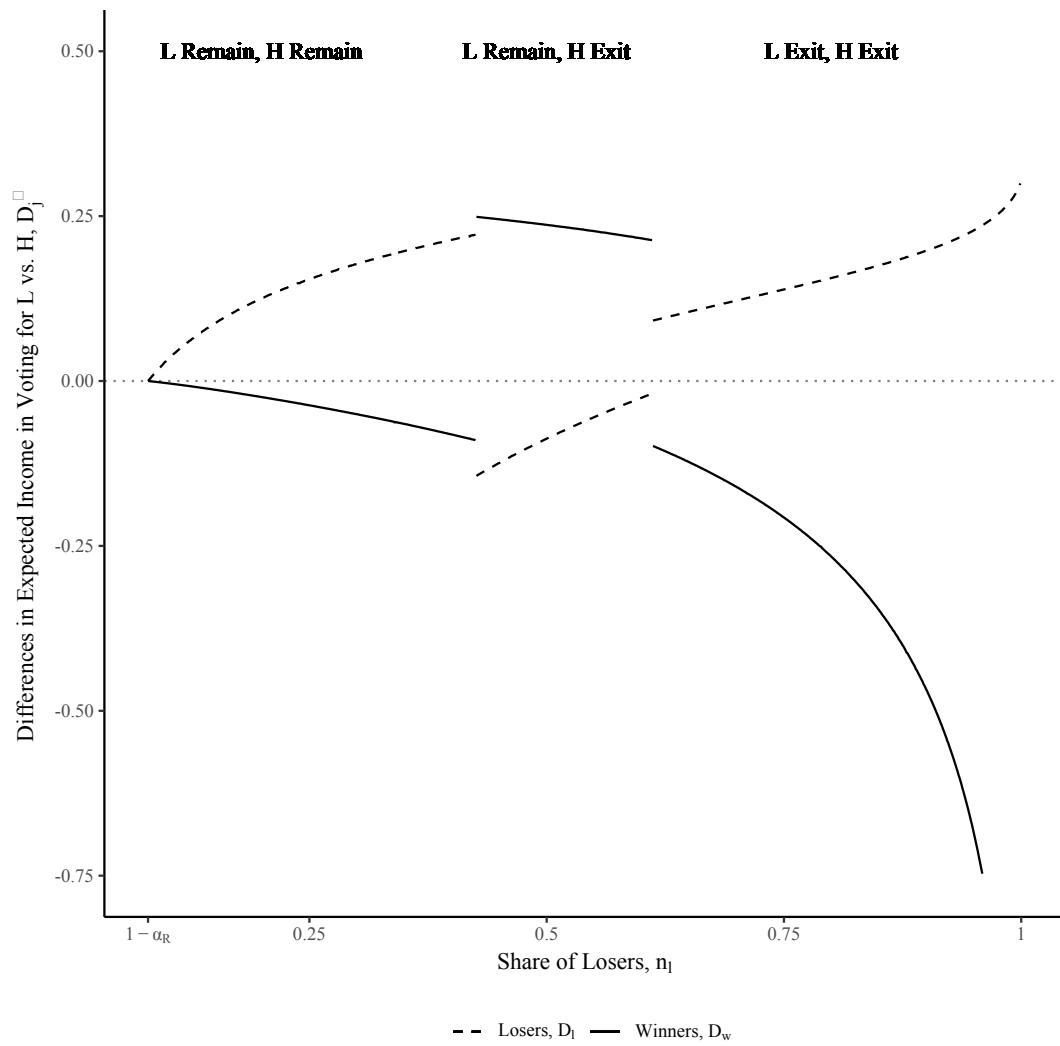


Figure 2: Inequality and Political Realignments

Political realignments can be seen in Figure 2, which plots D_w^* (solid line) and D_ℓ^* (dashed line), the differences in income that winners and losers expect when voting for L versus H as a function of the share of globalization losers. Each segment represents how D_w^* and D_ℓ^* change conditional on the policy outcomes that the candidates propose. Notice that D_ℓ^* is increasing in λ while D_w^* is decreasing in λ , representing direct, mechanical effects of increasing the share of individuals harmed by globalization who would subsequently prefer greater redistribution from winners to losers (cf. [Meltzer and Richard 1981](#)).

To see the intuition behind the proposition, consider how the sign of D_j^* changes at the discontinuities between the segments. In the leftmost case, where λ is low, both candidates would remain in the agreement. Political coalitions arise from the fact that H 's policymaking frictions limit his ability to redistribute away from α_R , leaving winners with greater income with H in power rather than L . L 's support comes from losers, to whom she can more easily redistribute. At the first discontinuity, the share of losers has grown such that H 's optimal action is now to exit rather than remain, which induces a reversal of political support. Now, winners coalesce behind L because they incur a large opportunity cost for abandoning the globalized status quo. This cost stems from three sources. If exit occurs, the winners' share of income prior to redistribution moves from α_R to α_E , decreasing their share of the economy. Exit is also inefficient, so winners are allocated a smaller share within a smaller pie. Additionally, the growing share of losers means that any redistribution becomes more costly, so winners expect greater taxation. By contrast, losers rally behind H , who, despite an inefficient foreign policy, ultimately apportions a greater share to them than L would. Finally, when λ is large, a second discontinuity implies that L too would now withdraw from the agreement. Political support switches again. Similar to the first case, winners support H because he can promise them a greater share of the new status quo distribution, $\theta_{HE}^* \geq \theta_{LE}^*$.

Inequality and Electoral Fortunes

In addition to changing the composition of political coalitions, rising inequality also directly affects politicians' electoral fortunes. Recall that the probability that L wins the election, $\pi(D_w, D_\ell)$ is determined by changes in group income as a function of the divisions of national income that L and H propose. In equilibrium, this is a piecewise function in three parts, corresponding to the three possible policy outcomes, which

can be written as

$$\pi^* = \begin{cases} \frac{1}{2} + b\left((1-\lambda)\log\left(\frac{\theta_{LR}^*}{\theta_{HR}^*}\right) + \lambda\log\left(\frac{1-\theta_{LR}^*}{1-\theta_{HR}^*}\right)\right) & \gamma < \gamma_H \\ \frac{1}{2} + b\left((1-\lambda)\log\left(\frac{\theta_{LR}^*}{\gamma\theta_{HE}^*}\right) + \lambda\log\left(\frac{1-\theta_{LR}^*}{\gamma(1-\theta_{HE}^*)}\right)\right) & \gamma_H \leq \gamma \leq \gamma_L \\ \frac{1}{2} + b\left((1-\lambda)\log\left(\frac{\theta_{LE}^*}{\theta_{HE}^*}\right) + \lambda\log\left(\frac{1-\theta_{LE}^*}{1-\theta_{HE}^*}\right)\right) & \gamma > \gamma_L. \end{cases}$$

It is difficult to make general, conclusive statements about how inequality affects electoral outcomes, but this exercise can shed light on the conditions under which ascendant, anti-globalization challengers pose electoral threats to more mainstream globalist candidates (Colantone and Stanig 2018; 2019). Increasing the share of losers in society has three distinct effects on the domestic political competition surrounding globalization policy. Consider the following decomposition of L 's probability of winning:

$$\frac{d\pi^*}{d\lambda} = \underbrace{\frac{\partial\pi^*}{\partial\lambda}}_{\text{direct effect}} + \underbrace{\frac{\partial\pi^*}{\partial\theta_{La}^*} \frac{\partial\theta_{La}^*}{\partial\lambda} + \frac{\partial\pi^*}{\partial\theta_{Ha}^*} \frac{\partial\theta_{Ha}^*}{\partial\lambda}}_{\text{strategic effect}}.$$

Clearly, there is a direct, mechanical effect of increasing the share of losers on the probability that L wins the election, since π is a population-weighted average of the incomes of winners and losers. The other two effects are strategic, and arise from shifts in candidates' redistributive proposals as a function of λ . We discuss each of these in turn, focusing on the substantively interesting case where L would remain in the agreement but H would exit. The following proposition previews the result.

Proposition 4 *Fix $\gamma_H \leq \gamma \leq \gamma_L$. If $\lambda \geq 1 - \theta_{HE}^*$ and $\lambda \leq 1 - \theta_{LR}^*$, then rising inequality adversely affects L 's electoral prospects, $\frac{d\pi^*}{d\lambda} \leq 0$.*

The direct and indirect effects of inequality on electoral fortunes need not move in the same direction, complicating our ability to make definitive statements. The direct effect concerns shifts in the composition of the population. In the case where $\gamma_H \leq \gamma \leq \gamma_L$, we know that losers would support H and winners would support L . Therefore, increasing the share of losers implies a greater voting bloc opposed to L , implying a negative direct effect. Moreover, increasing the share of losers also affects how candidates propose to divide national income. There are thus two relevant indirect effects, stemming from H 's proposal θ_{HE}^* and L 's proposal θ_{LR}^* . These indirect effects are at odds with one another, and bound the range of λ where inequality can harm L electorally. Recall that from Corollary 1, candidates must compensate more when the share of losers goes up, i.e., $\frac{\partial\theta_{La}^*}{\partial\lambda} < 0$. Therefore, for H 's rising redistributive proposal to have adverse electoral effects on L , λ must be bounded from below. In other words, there must be enough globalization

losers for changes in H 's policy proposal to affect L 's political support. By contract, for the indirect effect from L 's proposal to be disadvantageous to her, λ cannot be too high. Although L lacks support from the losers, she can redistribute more easily. In this case, L attempts to match H 's proposal to the losers as “quickly” as possible as inequality grows. Thus, in this range of values of λ , both indirect effects are negative, establishing a negative relationship between inequality and L 's electoral prospects.

Empirically, it has become more common that pro-globalization establishment parties lose out in elections to anti-globalization challengers (Walter 2021a). While the model can produce this result, the above analysis demonstrates that relatively stringent conditions must hold for a pro-globalization candidate's electoral chances to decrease with rising inequality. How then can we square the theoretical findings with observed data? One explanation might be that the shortcomings of redistributive policy in democratic polities are correlated with rising inequality, so the magnitude of the disadvantageous direct effect exceeds the potentially positive indirect effects. Hence, because democracies have in general failed to sufficiently compensate citizens (Bonica et al. 2013), the returns to compensation for globalization losers, even if positive, may be overwhelmed by the negative effects of increasing inequality within a globalized society.

Rising inequality unambiguously pressures the system of domestic policies upon which the survival of globalization is predicated. When embedded liberalism is needed most, its promises are not credible. When λ is low and neither L nor H would exit the agreement, winners support H who will redistribute less than L . This fact is consistent with the idea that winners are increasingly unwilling to share their gains from specialization with globalization losers (Linardi and Rudra 2020). However, a growing λ renders exit a credible option for H , who finds new political support from the losers. With alternative policies on the table, losers find compensation offered by pro-globalization L to be insufficient (Bowen, Broz and Rosendorff 2022). Subsequently, they prefer to support an anti-integrationist candidate instead. By contrast, winners now *ex post* would have preferred to maintain a system of embedded liberalism, but the demand for redistribution from the losers no longer exists. These tensions exacerbate further with continued inequality, because political candidates find it even more favorable to withdraw from the agreement rather than uphold it. Ultimately, concerns for domestic political survival trump international integration when inequality rises because leaders have incentives to sacrifice liberalization for political expedience.

Discussion

Central to the analysis is the notion that leaders vary in their costs of policymaking, or their ability to “get things done.” The policy frictions induced by these costs serve as a commitment device to credibly execute a welfare inferior policy like exit. High policymaking costs stifle the pursuit of domestic policies that are required in order to sustain globalization, like redistribution. This generates friction in a candidate’s ability to transfer income between winners and losers; consequently, high-cost candidates pursue other policy measures like exit to recalibrate the economic standing of winners and losers and their subsequent political loyalties.

When a candidate with high policymaking costs cannot credibly exit from an agreement, he finds political support with the globalization winners because his redistributive proposals are biased toward the status quo. This status quo bias can help to explain the variation in policy support and subsequent composition in the winning coalition of the Republican Party in the United States. Since the mid-twentieth century, Republicans have traditionally supported globalization winners and elites, preferring fewer tariffs and lower levels of redistribution than Democrats ([Irwin 2017](#)). By promoting views of globalization winners, Republicans upheld the status quo and continued the United States’ commitment to international integration, while simultaneously advancing domestic policies to cut rather than expand redistribution to the globalization losers. Such policies have increased inequality in large part due to globalization ([Ravallion 2018](#)). Moreover, places exposed most heavily to globalization saw some of the greatest decline in compensation ([Autor, Dorn and Hanson 2013](#)).

The growing number of globalization losers in the United States not only meant that the number of individuals requiring compensation increased, but it also shifted the willingness of Republican political candidates to continue to uphold pro-globalization policies. The rise of right-wing, anti-globalization candidates, even predating the presidency of Donald Trump, confirms this trend ([Cerrato, Ferrara and Ruggieri 2018](#); [Kuk, Seligsohn and Zhang 2022](#)). These candidates found it too costly to maintain a system of international liberalization and domestic redistribution: with rising inequality, many Republican candidates abandoned their support of globalization and ran instead on a platform promoting anti-globalization measures which include withdrawals from international agreements. Consequently, the locus of their political support transitioned from winners to losers, ushering a realignment both in supply and demand for pro-globalization policy in American politics across parties ([Schonfeld 2021](#)).

When initially laying the foundations for the contemporary system of international integration, it was assumed that distributional tensions would be assuaged through domestic policies of adjustment; governments

simply had to spend more to compensate globalization losers ([Cameron 1978](#); [Ruggie 1982](#); [Rodrik 1998](#)). However, embedded liberalism relies on a social compact that is *ex ante* undesirable for winners, but *ex post* insufficient for losers. Under the globalized status quo, losers' support for L establishes the demand for social programs to compensate those disaffected by globalization ([Walter 2010](#); [Rickard 2015](#)). Several analyses of the Trade Adjustment Assistance program in the United States have shown that exposure to compensation mitigates demand for protection, both in the form of submitting antidumping petitions ([Kim and Pelc 2021a](#)) and supporting Donald Trump's 2016 presidential campaign ([Ritchie and You 2021](#)). However, the emergence of anti-globalization candidates undermines the credibility of the demand for redistribution. [Milner \(2021\)](#) shows that once anti-globalization measures like exit become credible, even individuals who receive welfare benefits within areas shocked by international competition vote for anti-integrationist candidates, demonstrating how adjustment is no longer viewed as enough to compensate damages to losers (cf. [Bowen, Broz and Rosendorff 2022](#)). Conversely, winners find themselves wishing they had supplied requisite transfers to sustain liberalization: while winners initially supported candidates who would redistribute less, they would receive a smaller share of a smaller pie should exit occur.

References

- 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.
- Bonica, Adam, Nolan McCarty, Keith T. Poole and Howard Rosenthal. 2013. “Why Hasn’t Democracy Slowed Rising Inequality?” *Journal of Economic Perspectives* 27(3):103–124.
- Bornschieer, Simon. 2017. “Populist Mobilization Across Time and Space: An Introduction.” *Swiss Political Science Review* 23(4):301–312.
- 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.
- 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.
- Colantone, Italo and Piero Stanig. 2019. “The Surge of Economic Nationalism in Western Europe.” *Journal of Economic Perspectives* 33(4):128–151.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. Boston: Harper and Row.
- Egger, Peter H., Sergey Nigai and Nora M. Strecker. 2019. “The Taxing Deed of Globalization.” *American Economic Review* 109(2):353–390.
- Flaherty, Thomas M. and Ronald Rogowski. 2021. “Rising Inequality As a Threat to the Liberal International Order.” *International Organization* 75(2):495–523.
- Ge, Zoe Xincheng. 2022. “Obsolescent Treaties: Global Value Chains and the Termination of Bilateral Investment Treaties.”
- 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.
- Hays, Jude C. 2009. *Globalization and the New Politics of Embedded Liberalism*. Oxford: Oxford University 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.
- Kim, Sung Eun and Krzysztof J. Pelc. 2021a. “The Politics of Trade Adjustment Versus Trade Protection.” *Comparative Political Studies* 54(13):2354–2381.
- Kim, Sung Eun and Krzysztof Pelc. 2021b. “How responsive is Trade Adjustment Assistance?” *Political Science Research and Methods* 9(4):889–898.
- 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.
- Malis, Matt, B. Peter Rosendorff and Alastair Smith. 2022. “A Political Economy of International Organizations.”
- Mansfield, Edward D., Helen V. Milner and Nita Rudra. 2021. “The Globalization Backlash: Exploring New Perspectives.” *Comparative Political Studies* 54(13):2267–2285.
- Melnick, Justin and Alastair Smith. 2022. “International Negotiations in the Shadow of Elections.” *Journal of Conflict Resolution* .
- Meltzer, Allan H. and Scott F. Richard. 1981. “A Rational Theory of the Size of Government.” *Journal of Political Economy* 89(5):914–927.
- 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.
- Putnam, Robert D. 1988. “Diplomacy and Domestic Politics: The Logic of Two-Level Games.” *International Organization* 42(3):427–460.
- Ravallion, Martin. 2018. “Inequality and Globalization: A Review Essay.” *Journal of Economic Literature* 56(2):620–642.
- Rickard, Stephanie J. 2015. “Compensating the Losers: An Examination of Congressional Votes on Trade Adjustment Assistance.” *International Interactions* 41(1):46–60.
- Ritchie, Melinda N. and Hye Young You. 2021. “Trump and Trade: Protectionist Politics and Redistributive Policy.” *The Journal of Politics* 83(2):800–805.
- Rodrik, Dani. 1998. “Why do More Open Economies Have Bigger Governments?” *Journal of Political Economy* 106(5):997–1032.

- Rodrik, Dani. 2018. "Populism and the economics of globalization." *Journal of International Business Policy* 1(1):12–33.
- 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.
- Stolper, Wolfgang F. and Paul A. Samuelson. 1941. "Protection and Real Wages." *The Review of Economic Studies* 9(1):58–73.
- 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. 2010. "Globalization and the Welfare State: Testing the Microfoundations of the Compensation Hypothesis." *International Studies Quarterly* 54(2):403–426.
- 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

Candidates L and H choose to remain or exit from an international agreement, $\rho_d \in \{0, 1\}$ where $\rho_d = 1$ signifies remain for $d = L, H$. In addition to choosing whether to remain or exit, candidates choose redistributive policies $\theta_{da} \in [0, 1]$ for $a = \{remain, exit\}$. The share θ_{da} represents the amount of the economy allocated to winners. That is, candidates decide whether to stay or leave and how much they want to redistribute off of the status quo income distributions θ_a . A strategy for candidate d is therefore a choice $\rho_d \in \{0, 1\}$ and proposals $\theta_{da} \in [0, 1]^2$.

Voters observe candidate proposals and decide whether to vote for L or for H . Voters are either winners or losers, indexed by j . There are $1 - \lambda$ winners and λ losers for $\lambda \in [0, 1]$. Voters receive per capita payoffs based on the income allocated to their group by candidate d . Let voters have increasing and concave payoffs $v(\cdot)$ over income. For simplicity, we work with 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 $\log(\frac{\theta_{LR}}{1-\lambda})$ and would be $\log(\frac{1-\theta_{LR}}{\lambda})$ to losers. Payoffs are more thoroughly defined in Table 1 of the main text.

In addition to utility over income, voter i in group j receives a common shock $\beta \sim U[-\frac{1}{2b}, \frac{1}{2b}]$ and an individual shock $\mu_{ij} \sim U[-\frac{1}{2m_j}, \frac{1}{2m_j}]$ “in favor” of L . Therefore, given L ’s proposal θ_{La} and H ’s proposal θ_{Ha} , voter i votes for L if and only if $v_{ij}(\theta_{La}) + \beta + \mu_{ij} \geq v_{ij}(\theta_{Ha})$. A strategy for voter i in group j is a choice to vote for L or for H given the candidates’ choices to remain or exit and the accompanying proposals θ_{da} , as well as the values of their valence shocks. Formally, $\sigma_j : \{0, 1\}^2 \times [0, 1]^4 \times [-\frac{1}{2m}, \frac{1}{2m}] \times [-\frac{1}{2b}, \frac{1}{2b}] \rightarrow \{L, H\}$.

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

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

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, candidates optimally select redistributive proposals θ_{da} and whether to remain or exit the agreement.

Proofs

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

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

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

Differentiating with respect to θ_{HR} yields

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

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

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

To rank the shares, note that the cross partials $\frac{\partial^2 u}{\partial \theta_{HR} \partial \kappa} = \alpha_R - \theta_{HR} \geq 0$ and $\frac{\partial^2 u}{\partial \theta_{HR} \partial \alpha_R} = \kappa > 0$. Therefore by the implicit function theorem, $\frac{\partial \theta_{HR}^*}{\partial \kappa} \geq 0$ and $\frac{\partial \theta_{HR}^*}{\partial \alpha_R} > 0$. Since $\alpha_E \leq \alpha_R$, it must be that $\theta_{dE} \leq \theta_{dR}$. Further, since $\kappa > 1$, we have that $\theta_{La} < \theta_{Ha}$. Combining these yields $\theta_{LE}^* \leq \theta_{HE}^* \leq \theta_{LR}^* \leq \theta_{HR}^*$. ■

Proof of Corollary 1: Proposition 1 establishes that $\frac{\partial \theta_{da}^*}{\partial \kappa_d} \geq 0$ and $\frac{\partial \theta_{da}^*}{\partial \alpha_a} > 0$. The cross partial $\frac{\partial^2 u}{\partial \theta_{da} \partial \lambda} = -\frac{b\Psi}{(1 - \theta_{da})\theta_{da}} < 0$. Therefore by the implicit function theorem $\frac{\partial \theta_{da}^*}{\partial \lambda} < 0$. ■

Proof of Corollary 2: Recall $\eta(x) = \frac{\lambda x}{(1 - \lambda)(1 - x)}$ for some share of the economy x owned by the winners. Differentiating yields $\frac{\partial \eta(\theta_{da}^*)}{\partial \lambda} = \frac{\theta_{da}^*(1 - \theta_{da}^*) + (1 - \lambda)\lambda \frac{\partial \theta_{da}^*}{\partial \lambda}}{(1 - \lambda)^2(1 - \theta_{da}^*)^2}$. In the numerator, $\frac{\partial \theta_{da}^*}{\partial \lambda} \leq 0$ from Corollary 1. However, if $|\frac{\partial \theta_{da}^*}{\partial \lambda}| \leq \frac{\theta_{da}^*(1 - \theta_{da}^*)}{\lambda(1 - \lambda)}$, then $\frac{\partial \eta}{\partial \lambda} \geq 0$. From the implicit function theorem, $|\frac{\partial \theta_{da}^*}{\partial \lambda}| = \frac{b\Psi}{(1 - \theta_{da}^*)^2\theta_{da}^{*2}\kappa_d + b\Psi((1 - \theta_{da}^*)^2 - \lambda(1 - 2\theta_{da}^*))} \theta_{da}^*(1 - \theta_{da}^*)$. Clearly, $\frac{b\Psi}{(1 - \theta_{da}^*)^2\theta_{da}^{*2}\kappa_d + b\Psi((1 - \theta_{da}^*)^2 - \lambda(1 - 2\theta_{da}^*))} \leq \frac{1}{\lambda(1 - \lambda)}$, so $\frac{\partial \eta(\theta_{da}^*)}{\partial \lambda} \geq 0$. ■

Proof of Proposition 2: We will establish existence of γ_L , the proof for γ_H is analogous. Let p_d be the

probability that candidate d exits. L 's indirect utility functions for remaining and exiting are

$$\begin{aligned} EU_L(\text{remain}) &= \frac{1}{2}(\Psi - (\alpha_R - \theta_{LR}^*)^2) + b\Psi \left(p_H \lambda \log\left(\frac{\theta_{HE}^*(1 - \theta_{HR}^*)}{\theta_{HR}^*(1 - \theta_{HE}^*)}\right) - p_H \log\left(\frac{\gamma \theta_{HE}^*}{\theta_{HR}^*}\right) \right. \\ &\quad \left. + \log\left(\frac{\theta_{LR}^*}{\theta_{HR}^*}\right) + \lambda \log\left(\frac{\theta_{HR}^*(1 - \theta_{LR}^*)}{\theta_{LR}^*(1 - \theta_{HR}^*)}\right) \right). \\ EU_L(\text{exit}) &= \frac{1}{2}(\Psi - (\alpha_E - \theta_{LE}^*)^2) + b\Psi \left(p_H \lambda \log\left(\frac{\theta_{HE}^*(1 - \theta_{HR}^*)}{\theta_{HR}^*(1 - \theta_{HE}^*)}\right) - p_H \log\left(\frac{\gamma \theta_{HE}^*}{\theta_{HR}^*}\right) \right. \\ &\quad \left. + \log\left(\frac{\gamma \theta_{LE}^*}{\theta_{HR}^*}\right) + \lambda \log\left(\frac{\theta_{HR}^*(1 - \theta_{LE}^*)}{(1 - \theta_{HR}^*)\theta_{LE}^*}\right) \right). \end{aligned}$$

Therefore, L prefers to remain whenever

$$\Phi_L(\gamma) := \frac{1}{2}(\alpha_E + \alpha_R - \theta_{LE}^* - \theta_{LR}^*)(\alpha_E - \alpha_R - \theta_{LE}^* + \theta_{LR}^*) + b\Psi \left(\log\left(\frac{\theta_{LR}^*}{\gamma \theta_{LE}^*}\right) + \lambda \log\left(\frac{\theta_{LE}^*(1 - \theta_{LR}^*)}{\theta_{LR}^*(1 - \theta_{LE}^*)}\right) \right) > 0.$$

Since $\Phi_d(\gamma)$ is strictly decreasing in γ , $\frac{\partial \Phi_d}{\partial \gamma} = -\frac{b\Psi}{\gamma}$, by the intermediate value theorem there is a point γ_d where $EU_d(\text{remain}) = EU_d(\text{exit})$ with $EU_d(\text{exit}) > EU_d(\text{remain})$ whenever $\gamma > \gamma_L$.

To rank, we want to demonstrate that $\frac{\partial \gamma_H}{\partial \kappa} \leq 0$. Because κ_L is normalized to 1, we know that increasing κ above 1 means that exit is more likely. We know that γ_H solves $\Phi_H(\gamma) = EU_H(\text{remain}) - EU_H(\text{exit}) = 0$ at equality, analogous to above. Differentiating yields $\frac{\partial \Phi_H}{\partial \kappa} = \frac{1}{2}(\alpha_E^2 - \alpha_R^2 - 2\alpha_E \theta_{HE}^* + \theta_{HE}^{*2} + 2\alpha_R \theta_{HR}^* - \theta_{HR}^{*2}) \leq 0$. Therefore by the implicit function theorem, $\frac{\partial \gamma_H}{\partial \kappa} \leq 0$. ■

Proof of Corollary 3: We want to compute $\frac{\partial \gamma_d}{\partial \lambda} = -\frac{\partial \Phi_d / \partial \lambda}{\partial \Phi_d / \partial \gamma}$. From Proposition 2, $\frac{\partial \Phi_d}{\partial \gamma} < 0$. By the envelope theorem, the first term of $\frac{d \Phi_d}{d \lambda} = \frac{\partial \Phi_d}{\partial \theta_{da}} \frac{\partial \theta_{da}}{\partial \lambda} + \frac{\partial \Phi_d}{\partial \lambda}$ is zero. Differentiating, $\frac{\partial \Phi_d}{\partial \lambda} = b\Psi \log\left(\frac{\theta_{dE}^*(1 - \theta_{dR}^*)}{\theta_{dR}^*(1 - \theta_{dE}^*)}\right) < 0$ because $\theta_{dE}^* \leq \theta_{dR}^*$ by Proposition 1. Then, by the implicit function theorem, $\frac{\partial \gamma_d}{\partial \lambda} \leq 0$. ■

Proof of Proposition 3: It is sufficient to demonstrate that there exist cases where increasing λ such that the equilibrium outcome moves from both L and H remaining to L remaining and H exiting implies that D_w^* and D_ℓ^* change sign. Consider, for instance, a case where $\alpha_R = 0.92$, $\theta_E = 0.6$, $b = 1$, $\Psi = 1$, and $\kappa = 2.5$ as primitives. Let $\lambda = 0.425$. Equilibrium shares are then $\theta_{LR}^* = 0.68$, $\theta_{LE}^* = 0.58$, $\theta_{HE}^* = 0.59$, and $\theta_{HR}^* = 0.74$. With these shares, we calculate $\gamma_H = 0.9002$ and $\gamma_L = 0.95$. Pick $\gamma = 0.9$ so that neither H nor L can credibly exit. Then $D_w^* = -0.09$ so winners support H and $D_\ell^* = 0.22$ so losers support L .

Now increase $\lambda = 0.43$. This leads to new proposals $\theta_{LR}^* = 0.677$, $\theta_{LE}^* = 0.58$, $\theta_{HR}^* = 0.74$, and $\theta_{HE}^* = 0.586$. With these shares, $\gamma_H = 0.898$ and $\gamma_L = 0.947$, so H prefers to exit but L does not. Then $D_w^* = 0.25$, so winners support L , and $D_\ell^* = -0.14$, so losers support H , establishing existence. ■

Proof of Proposition 4: If $\gamma_H \leq \gamma \leq \gamma_L$, the probability that L wins the election is

$$\pi^* = \frac{1}{2} + b \left((1 - \lambda) \log\left(\frac{\theta_{LR}^*}{\gamma \theta_{HE}^*}\right) + \lambda \log\left(\frac{1 - \theta_{LR}^*}{\gamma(1 - \theta_{HE}^*)}\right) \right).$$

Differentiating yields

$$\frac{d\pi^*}{d\lambda} = b \left(\log \left(\frac{(1 - \theta_{LR}^*)\theta_{HE}^*}{(1 - \theta_{HE}^*)\theta_{LR}^*} \right) - \frac{1 - \lambda - \theta_{HE}^*}{(1 - \theta_{HE}^*)\theta_{HE}^*} \frac{\partial \theta_{HE}^*}{\partial \lambda} + \frac{1 - \lambda - \theta_{LR}^*}{(1 - \theta_{LR}^*)\theta_{LR}^*} \frac{\partial \theta_{LR}^*}{\partial \lambda} \right).$$

The first term is the direct effect and is always negative. The next two terms represent parts of the indirect effect owing to H and L 's policy choices. A sufficient condition for $\frac{d\pi^*}{d\lambda} < 0$ would be for both indirect effects to be negative. This occurs when $1 - \lambda - \theta_{HE}^* \leq 0$ for the second term and $1 - \lambda - \theta_{LR}^* \geq 0$ for the third term. Therefore, when $\lambda \in [1 - \theta_{HE}^*, 1 - \theta_{LR}^*]$ we have $\frac{d\pi^*}{d\lambda} < 0$. ■