Foreign Policy Analysis (2024), orae006

Top-Category Inflation in Ordered International Relations Outcomes

Benjamin E. Bagozzi
University of Delaware, USA
Minnie M. Joo
University of Massachusetts Lowell, USA
and
Bumba Mukherjee
Penn State University, USA

Ordered dependent variables are widely employed in international relations (IR). These ordered dependent variables often suffer from inflated observations in their highest outcome category due to distinct processes. The application of standard ordered probit and ordered logit models to such ordinal measures will fail to capture these distinct processes, often producing biased inferences as a result. Yet IR researchers have thus far ignored the potential for top-category inflation in ordered outcome variables. We sensitize researchers to top-category inflation in ordered IR outcomes. We then intuitively extend the widely used zero-inflated ordered probit model to the top-category inflated setting, providing resources to facilitate the proper modeling of top-category inflation in ordered outcomes. Finally, we provide two applications to published IR research related to trade politics and political repression. Together, these applications illustrate the substantive and methodological potentials of our proposed tools for diagnosing and modeling top-category inflation in IR outcomes.

Las variables dependientes ordenadas se utilizan de manera frecuente en el campo de las Relaciones Internacionales (RRII). Estas variables dependientes ordenadas suelen contener observaciones infladas en su categoría de resultado más alta debido a procesos distintos. La aplicación de los modelos *probit* ordenado estándar y *logit* ordenado sobre tales medidas ordinales no logra capturar estos procesos distintos, lo que a menudo produce inferencias sesgadas como resultado. Sin embargo, los investigadores en el campo de las RRII han ignorado, hasta ahora, el potencial de inflado de la categoría superior en las variables de resultado ordenadas. Sensibilizamos a los investigadores sobre el inflado de la categoría superior en los resultados ordenados del campo de las RRII. A continuación, extendemos intuitivamente el modelo *probit* ordenado inflado a cero, el cual es ampliamente utilizado, a la configuración inflada de la categoría superior, proporcionando así recursos para facilitar el modelado adecuado del inflado de la categoría superior en los resultados ordenados. Por último, proporcionamos dos aplicaciones para la investigación publicada en materia de las RRII, relacionadas con la política comercial y con la represión

Benjamin E. Bagozzi is an Associate Professor in the Department of Political Science and International Relations at the University of Delaware. His primary areas of specialization are political methodology, political violence, environmental politics, and international relations.

Minnie M. Joo is an Assistant Professor in the Department of Political Science at the University of Massachusetts Lowell. Her primary areas of research are civil conflict, peace processes, international relations, and political methodology.

Bumba Mukherjee is a Professor in the Department of Political Science at Penn State University. His primary areas of research are international political economy, civil conflict and political violence, and political methodology.

Bagozzi, Benjamin E. et al. (2024) Top-Category Inflation in Ordered International Relations Outcomes. *Foreign Policy Analysis*, https://doi.org/10.1093/fpa/orae006

política. En conjunto, estas aplicaciones ilustran el potencial sustantivo y metodológico de las herramientas propuestas para poder diagnosticar y modelar el inflado de la categoría superior en los resultados del campo de las RRII.

Les variables dépendantes classées sont largement employées en relations internationales (RI). Ces variables dépendantes classées pâtissent souvent d'observations gonflées quant à leur catégorie d'issues la plus élevée à cause de différents processus. L'application de modèles logit et probit classés standard à de telles mesures ordinales ne représente pas ces différents processus, et produit donc souvent des inférences biaisées. Pourtant, les chercheurs en RI ont jusqu'ici omis le potentiel de l'inflation de la catégorie supérieure dans les variables d'issues classées. Nous sensibilisons les chercheurs à l'inflation de la catégorie supérieure dans les issues de RI classées. Ensuite, nous suivons notre intuition en appliquant un modèle largement utilisé, le modèle probit classé à inflation de zéro, au cadre gonflé de la catégorie supérieure. Nous fournissons ainsi des ressources pour faciliter la modélisation en bonne et due forme de l'inflation de la catégorie supérieure dans les issues classées. Enfin, nous proposons deux applications aux travaux de recherche en RI publiés qui ont un lien avec la politique commerciale et la répression politique. La somme de ces applications illustre les potentiels méthodologiques importants des outils que nous proposons pour le diagnostic et la modélisation de l'inflation de la catégorie supérieure des issues en RI.

Limited dependent variables in international relations (IR) research—including binary variables, counts, ordered outcomes, duration measures, and unordered polytomous outcomes—have now been widely shown to be sensitive to problems of "zero-inflation." Zero-inflation can be defined as any instance where a variable's zero category contains a preponderance of observations arising from a distinct data-generating process (d.g.p.) from that producing the variable's remaining zero and nonzero observations. Ignoring this phenomenon not only leads to biased estimates and inferences, but also forgoes an opportunity to develop and test more nuanced theories of IR processes.

Thankfully, IR research now widely employs zero-inflated estimators to overcome these limitations. This includes efforts to account for zero-inflation in IR count measures as varied as human rights organizations' international advocacy activities (Murdie 2014), terrorist incidents across countries (Savun and Phillips 2009), and international country mentions in US Presidential Daily Briefings (Lebovic 2021). Scholars are also increasingly implementing these techniques to address similar zero-inflated processes in (i) binary dependent variables, such as interstate conflict (Xiang 2010), to disentangle politically relevant dyads in interstate conflict data, and (ii) ordered outcome variables, including among "peace" observations within studies of interstate and intrastate conflict (Bagozzi et al. 2015) and underreporting in measures of wartime sexual violence (Ju 2023).

Yet many ordinal IR measures do not exhibit signs of zero inflation and continue to be analyzed with standard ordered probit (OP) or ordered logit (OL) models. This includes ordered dependent variables encompassing processes of state repression, trade disputes, ex-combatants' trust in government institutions, respect for individual human rights, escape clauses in international trade agreements, and ratification of the Rome Statute of the International Criminal Court (ICC) (Piazza and Walsh 2009; Brysk and Mehta 2014; Baccini, Dür and Elsig 2015; Girod, Stewart, and Walters 2016). IR articles analyzing these ordered dependent variables encompass areas as diverse as postconflict peace, international political economy (IPE), international organizations, international law, and human rights. Our core contention

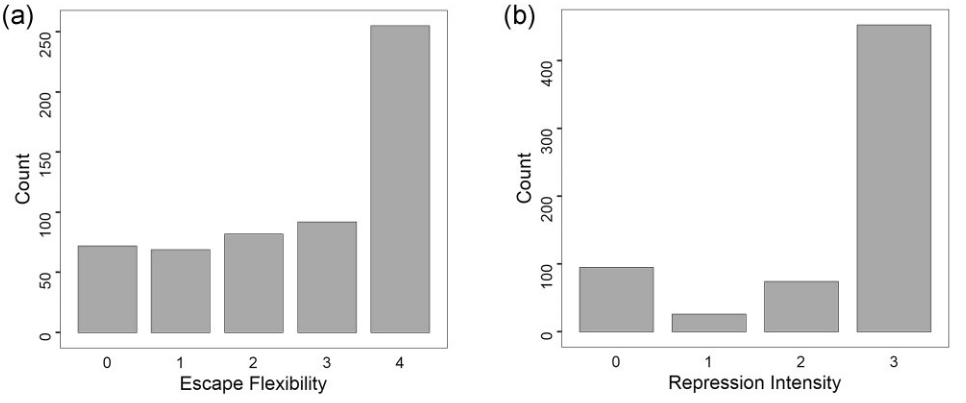

Figure 1. Top-category inflation

is that, in contrast to zero-inflation, many of these ordered IR dependent variables include inflated observations in the variable's *highest* (i.e., "top-category")—as opposed to the lowest—ordered outcome category. This top-category inflation can be accordingly defined as any instance where the top-category of an ordered variable contains a preponderance of observations arising from a distinct d.g.p. from that producing the variable's remaining (non)top-category observations.

Consider, for example, the frequency distribution of two ordered dependent variables analyzed later in this paper. These distributions appear in figure 1 and are drawn from Baccini, Dür, and Elsig's (2015) ordered measure of *Escape Flexibility* provisions in preferential trading agreements (PTAs)—which ranges from 0 (none) to 4 (maximum provisions)—and Girod, Stewart, and Walters' (2016) ordinal measure of *Repression Intensity*—which ranges from 0 (none) to 3 (extreme repression). The top-category of these ordered dependent variables includes a substantially higher share of observations compared to the other ordered outcome categories (figure 1a and b), rendering them "top-category inflated." These examples are hardly unique. Indeed, the top-category of the following additional ordered dependent IR variables in articles published in IO, ISQ, JCR, and JPR that are described, cited, and illustrated in the online appendix, figures A5, A7, and A9–A14, *also* have an excessive share of observations: state-perpetrated *Disappearances* following terrorist attacks, *ICC Ratification* by member states, state restrictions on *Religious Freedom*, *Trade concessions* by the World Trade Organization (WTO) member-states, interstate *War Outcome*, *economic threat* perceptions, ex-rebel combatants' *Trust in Reintegration* programs, and *Election Quality* monitoring.

We contend that excessive top-category observations in the aforementioned ordered dependent variables result from competing theoretical claims about state decision-making or behavior—i.e., distinct d.g.p.'s—that produce the *same* top-category outcome in such variables. Consequently, the top-category of ordered dependent variables that have a preponderant share of observations may include a mixture of (i) noninflated cases that emerge from the theoretically assumed ordered d.g.p. and its determinants *and* (ii) inflated cases that result from secondary processes that are not based on the theoretically assumed ordered d.g.p. and are hence better seen as "not applicable" (NA) cases. To see this more clearly, recall the excessive observations in *Repression Intensity*'*s* top-category of extreme repression. These excessive top-category observations include noninflated cases that are based on Girod et al.'s (2016) theoretically determined ordered continuum of incumbents unleashing substantial repression *targeted* against anti-government campaign-specific opposition activity. But it also includes inflated cases not based on the assumed ordered d.g.p. since these are cases, as described later, of incumbents employing extreme repression *indiscriminately* against noncampaign-specific personnel, including bystanders (Kalyvas 2006; Lyall 2009).

Likewise, the excessive observations in *Escape Flexibility*'s top-category of maximum PTA-flexibility provisions incorporate noninflated cases that, as per prevailing theories, are PTA member-states that negotiate and adopt flexibility provisions in a step-by-step manner to *insure* their economies against trade shocks while genuinely signaling their commitment to implement PTA-mandated trade reforms once the shock dissipates (Pelc 2009). It also includes inflated cases of PTA member-states that pursue maximum flexibility provisions concurrently for immediate protection of domestic special interests (Bhagwati et al. 1999). We further argue in the online appendix that the excessive top-category observations in each additional IR-ordered dependent variable listed earlier, including *Disappearances* and *ICC Ratification*, incorporate noninflated and inflated cases produced by distinct d.g.p.'s.

Apart from sensitizing researchers to the processes that lead to excessive top-category observations within ordered dependent variables, we demonstrate below that standard OP and OL models produce biased inferences in these contexts. To address this, we build on Harris and Zhao's (2007) zero-inflated OP approach and leverage Bagozzi et al.'s (2015) notation to develop a split-population model known as the top-category inflated ordered probit (TiOP) model, which can also be estimated with correlated errors. Hereafter, we refer to the latter model as the TiOPC model and refer to both TiOP and TiOPC models as the TiOP(C) models. While the TiOP estimator builds on the zero-inflated ordered probit (ZiOP) framework, we discuss in the following section why the TiOP(C) models have several advantages over the ZiOP(C) models when analyzing ordered dependent variables in which the top-category contains excessive observations. Further, unlike OP and OL models, TiOP(C) models jointly estimate two latent equations to statistically account for inflated top-category observations in ordered dependent variables: a probit inflation-stage equation and an augmented OP outcome equation. We explain in the next section how these features of the TiOP(C) model not only address key methodological limitations of OP and OL models in these contexts but also have important substantive implications for theoretical research in IR.

To preview, our Monte Carlo simulations and empirical applications reveal that, in contrast to the OP and OL models, the TiOP(C) estimators provide more accurate inferences, reduce chances of model misspecification, and detect nonmonotonic covariates effects when the ordered dependent variable's top-category observations include noninflated and inflated cases. Next, unlike the ZiOP(C) models, the TiOP(C) estimators avoid mischaracterizing top-category inflation processes in the context of the preponderant share of top-category observations in ordered dependent variables. Further, the TiOP(C) models also provide an opportunity for theoretical development given that they evaluate distinct theoretical processes that generate the observed top-category outcomes and analyze the effect of observables in the corresponding probit-inflation stage and OP-outcome stages.

We illustrate below the TiOP(C) model's advantages for addressing methodological challenges and theoretical development by estimating these models on published datasets that include the *Escape Flexibility* and *Repression Intensity* top-category inflated ordered dependent variables discussed above. These applications reveal that the TiOP(C) models not only yield rich substantive insights about the heterogeneous observations in the highest outcome category of inflated top-category ordered dependent variables, but also improve inferences about the effect of covariates on such variables. For instance, contrary to some key theoretical claims, the TiOP(C) models applied to *Escape Flexibility* indicate that (i) fully democratic PTA member-states pursue maximum flexibility provisions in PTAs to engage in import discrimination rather than for insurance against trade shocks, and (ii) nondemocratic PTA member-states are *not* associated with maximum flexibility provisions. The TiOP(C) models applied to *Repression Intensity* suggest that extreme repression in autocracies emerges from indiscriminate repressive behavior rather than from targeting specific anti-regime campaign activity, which challenges key claims in this issue area.

The TiOP(C) Statistical Framework

TiOP(C) Model(s)

The TiOP model accounts for top-category inflation in an ordered dependent variable y_i with $j = 0, 1, 2...J$ categories by combining two latent equations: (i) a probit inflation-stage equation that estimates the covariates' effect on the probability with which units in the top-category come from the noninflated versus inflated d.g.p. and (ii) an OP outcome equation that estimates how another set of covariates influences the probability of observing *each* category of the ordinal dependent variable, conditional on units being in the noninflated group. The errors from these two latent equations can be independent or correlated.

To define the TiOP(C) models, let $i \in 1, 2, 3...N$ denote each unit in the data. Let the binary variable s_i indicate the following split between units observed in y_i' s top (J^{th}) category: those in (i) regime 0 ($s_i = 0$), which are "inflated" cases whose coded top-category value in the said ordinal scale is inaccurate or determined by secondary theoretical processes that override the assumed ordinal continuum, and (ii) regime 1 ($s_i = 1$), which are noninflated cases whose coded top-category value in y_i is accurate. s_i is related to the latent dependent variable s_i^*, where $s_i = 1$ for $s_i^* > 0$; $s_i = 0$ for $s_i^* \leq 0$. The propensity with which units enter regime 1 is given by the probit inflation-stage equation:

$$s_i^* = \mathbf{z}_i'\gamma + u_i, \tag{1}$$

where $\mathbf{z}_i'$ is the covariate vector, γ the coefficient vector, and u_i the standard-normal error term. The probability of i being in regime 1 is $\Pr(s_i = 1|\mathbf{z}_i) = \Pr(s_i^* > 0|\mathbf{z}_i) = \Phi(\mathbf{z}'_i\gamma)$ and regime 0 is $\Pr(s_i = 0|\mathbf{z}_i) = \Pr(s_i^* \leq 0|\mathbf{z}_i) = 1 - \Phi(\mathbf{z}'_i\gamma)$. Φ is the standard-normal c.d.f. The TiOP(C) models' OP outcome-stage equation is

$$\widetilde{y}_i^* = \mathbf{x}_i'\beta + \varepsilon_i, \tag{2}$$

$$\widetilde{y}_i = \begin{cases} 0 & if \quad \widetilde{y}_i^* \leq 0 \\ j & if \quad \mu_{j-1} < \widetilde{y}_i^* \leq \mu_j \ (j = 1, ..., J-1) \\ J & if \quad \mu_{j-1} \leq \widetilde{y}_i^* \end{cases}, \tag{3}$$

where $\mathbf{x}'_i$ is the covariates vector, β the coefficient vector, and ε_i the standard-normal error term, $j = 1, 2..., J-1$ are observed values on $\widetilde{y}_i$, and μ_j (where $\mu_{j=0} = 0$ are the cut-point parameters). If the probit inflation stage and OP equation errors (u_i, ε_i) are *not* correlated, then the TiOP model's OP outcome equation *without* correlated errors is

$$\Pr(y_i) = \begin{cases} \Pr(y_i = 0|\mathbf{x}_i, \mathbf{z}_i) = [\Phi(\mathbf{z}_i'\gamma)\Phi(-\mathbf{x}_i'\beta)] \\ \Pr(y_i = j|\mathbf{x}_i, \mathbf{z}_i) = \Phi(\mathbf{z}_i'\gamma)[\Phi(\mu_j - \mathbf{x}_i'\beta) - \Phi(\mu_{j-1} - \mathbf{x}_i'\beta))] \\ \qquad (j = 1, ..., J-1) \\ \Pr(y_i = J|\mathbf{x}_i, \mathbf{z}_i) = [1 - \Phi(\mathbf{z}_i'\gamma)] + \Phi(\mathbf{z}_i'\gamma)[1 - \Phi(\mu_{j-1} - \mathbf{x}_i'\beta)] \end{cases}. \tag{4}$$

When u_i and ε_i are correlated and follow a bivariate-normal distribution with correlation coefficient $\rho_{\varepsilon u}$, the TiOPC model's OP outcome equation *with* correlated errors is

$$\Pr(y_i) = \begin{cases} \Pr(y_i = 0|\mathbf{x}_i, \mathbf{z}_i) = [\Phi_2(\mathbf{z}_i'\gamma, -\mathbf{x}_i'\beta; -\rho_{\varepsilon u})] \\ \Pr(y_i = j|\mathbf{x}_i, \mathbf{z}_i) = \Phi_2(\mathbf{z}_i'\gamma, \mu_j - \mathbf{x}_i'\beta; -\rho_{\varepsilon u}) \\ \qquad\qquad\qquad\qquad -\Phi_2(\mathbf{z}_i'\gamma, \mu_{j-1} - \mathbf{x}_i'\beta; -\rho_{\varepsilon u}) \\ \Pr(y_i = J|\mathbf{x}_i, \mathbf{z}_i) = [1 - \Phi(\mathbf{z}_i'\gamma)] + \Phi_2(\mathbf{z}_i'\gamma, \mathbf{x}_i'\beta - \mu_{j-1}; \rho_{\varepsilon u}) \end{cases}. \quad (5)$$

Φ_2 denotes the bivariate normal distribution's c.d.f. Equation (1) constitutes the TiOP and TiOPC model's first (inflation) stage. The *augmented* OP outcome equations—which constitute the TiOP and TiOPC models' combined first and second stages—are given by Equations (4) and (5), respectively. The TiOP(C) models, whose (log-)likelihoods are defined in the online appendix, jointly estimate the inflation-stage probit equation and their augmented OP-stage equations.

Note that the probability of observing the top-category in the TiOP(C) models' augmented OP equation is modeled conditional upon the probability of an observation arising from the ordered d.g.p. *plus* the probability of that observation being an inflated case. This feature statistically accounts for inflated and noninflated cases in the ordered dependent variable's (y_i) top-category that result from two d.g.p.'s and illuminates *when* the inflated top-category cases result from secondary theoretical processes. Accordingly, our Monte Carlo experiments that evaluate the performance of the TiOP, TiOPC, and OP models and diagnostic tests for top-category inflation—presented in online appendix, figures A15–A20 and tables A8–A13—reveal that the TiOP(C) models outperform the OP model under TiOP(C) d.g.p.'s.

Methodological and Substantive Contributions

The TiOP(C) estimators address several methodological limitations of the OP, OL, and ZiOP(C) models and have important implications for theoretical research when analyzing ordered dependent variables with top-category inflation. Specifically, from a methodological perspective, IR scholars often code the top-category of ordered dependent variables as a summation category for multiple lower-category pathways. This is likely to be especially acute for "less granular" ordered variables and, when present, will produce excessive top-category observations, which include heterogeneous inflated and noninflated cases that arise from distinct theoretical mechanisms. Standard OP and OL models treat such heterogeneous cases as homogeneous by default since these models cannot account for the preponderant share of top-category observations. This leads to inaccurate inferences. The TiOP(C) models' two latent equations, however, explicitly account for the heterogeneous set of inflated and noninflated top-category observations, which facilitates accurate inferences.

Another methodological limitation of OP and OL models is that they neither assess nor account for the explanatory factors that generate the mixture of inflated and noninflated observations in inflated ordinal dependent variables. This leads to model misspecification. The TiOP(C) estimators avoid such misspecification as they accommodate theoretically identified variables in the models' probit inflation-stage *and* augmented OP outcome equations. Consequently, the TiOPC(C) models do not over(under)estimate the effect of covariates due to this issue and can detect nonmonotonic covariate effects.

Unlike the OP and OL models, researchers can technically use the ZiOP model to address top-category inflation in an ordered dependent variable by manually reverse-coding their top-category inflated ordered variable prior to analysis. Yet—as evidenced by the scarcity of extant research utilizing such a work-around—this solution has its limitations. First, reverse coding many ordered variables will often lead to counterintuitive—or laborious—interpretations of coefficient estimates and marginal effects in relation to one's theoretical process of interest. While not an insurmountable challenge in and of itself, adding such opacity to the (inflation or outcome stage) contingent interpretations of multiequation-inflated models will often undermine a researcher's ability to convey empirical findings effectively and efficiently to the reader. Moreover, the absence of easy-to-use R code for ZiOP models further limits the usefulness of a reverse coding work-around. Our TiOP(C) models address these challenges as they do not require any reverse coding of one's dependent variable. We also provide R code for its implementation.

Finally, the TiOP(C) estimators are more than just a methodological fix. They also have three key substantive implications for theoretical research. First, the TiOP(C) models permit researchers to investigate the nature of the theoretical processes causing observed top-category outcomes and to analyze the varying effects of observables in the probit-inflation stage and OP-outcome stage. This is crucial because IR scholars do not theoretically account for the possibility that top-category observations in ordered dependent variables for outcomes such as state repression, war results, and treaty ratification (among others) may incorporate heterogeneous units whose distinct behavior can produce the same top-category outcome. Accordingly, the TiOP(C) models provide an opportunity for IR scholars to develop and evaluate more nuanced yet comprehensive theories that fully account for the political outcomes that noninflated and inflated cases exhibit in the top-category of ordered dependent variables that have excessive observations.

Second, the TiOP(C) models' two latent equations can capture differences in decision-making by different units (e.g., states) that produce the same inflated top-category outcome. This enables researchers to closely tie their statistical model to theory and vice versa, which is vital for empirical inference dating back to the inception of the EITM (Empirical Implications of Theoretical Models). It also provides room to construct and test richer theoretical micro-foundations about which states, governments, or other units may—and which may not—enter the ordered dependent variable's inflated top-category outcome. Third, the TiOP(C) estimators account for unmodeled factors that may influence theoretical processes in the probit-inflation and OP-outcome stages. This not only recovers useful information that otherwise would be lost but also makes it imperative for scholars to develop more fine-grained theories to unpack latent factors that may provide greater explanatory leverage to account for a wide variety of IR-ordered outcomes.

Applications

We apply our TiOP(C) models to datasets from two published works in IR that use conventional OP (or OL) models: Baccini et al.'s (2015) research on "Escape Flexibility" provisions in PTAs analyzed by IPE scholars, and Girod et al.'s (2016) work on state repression that speaks to the human rights literature. Following studies on how escape clauses facilitate compliance with international agreements (Rosendorff and Milner 2001; Pelc 2009), Baccini et al. develop a dataset of 587 PTAs during 1945–2009 to analyze when PTA member-states negotiate to incorporate four Escape Flexibility provisions: safeguards (SGs), suspension of tariff cuts (STCs), anti-dumping duties (ADs), and countervailing duties (CVDs). They thus operationalize *Escape Flexibility* as a 0-to-4 ordered dependent variable in which the ordered continuum directly captures their theoretical claims about how and when escape flexibility provisions are adopted by PTA member-states. *Escape Flexibility* is coded as 0 when a PTA does not include any flexibility provision, 1 for PTAs that only include the minimum (i.e., one) flexibility provision of SGs, 2 for SGs and STCs, 3 for three flexibility provisions (SGs, ADs, and STCs), and 4 for maximum (all four flexibility) provisions: SGs, STCs, ADs, and CVDs.

Baccini et al. employ a standard OP model to test whether: (i) greater *Depth* of trade liberalization commitments in PTAs increase *Escape Flexibility* (Hypothesis 1),

(ii) *Depth*'s positive influence on *Escape Flexibility* is weaker for democracies than for nondemocracies (Hypothesis 2), using the *Depth* × *Regime Dummy* interaction term where *Regime Dummy* denotes fully democratic PTA member-states,[1] and (iii) PTA-member countries that are fully democratic and those that have recently experienced *Democratization* favor maximum provisions.[2] They control for *Trade* flows, *GDP*, *GDPpc* per PTA, a *WTO* dummy for PTA member-states in the WTO, and the number of PTA member-states (*No. Members*).[3]

Baccini et al.'s main OP specification results in Model 3 (online appendix, table A3) that we focus on for our TiOP(C) application support Hypothesis 1 stated above. The marginal effect of *Depth* × *Regime Dummy* that they illustrate statistically corroborates Hypothesis 2's prediction that nondemocracies but *not* democracies are significantly more likely to include maximum flexibility provisions in PTAs (p. 773). They also find that democracies incorporate maximum PTA-flexibility provisions *irrespective* of the level of trade reform commitments (*Depth*), while *Democratization* is negative and significant. All controls are insignificant, barring *Trade*.

Baccini et al.'s results provide rich insights. But their standard OP models overlook the fact that the ordered *Escape Flexibility*'s top-category of maximum (all four) flexibility provisions incorporates three times more observations than any of its other categories (figure 1a). Their OP model estimates are also potentially biased as they do not account for the possibility that the excessive top-category (maximum flexibility provision) observations in *Escape Flexibility* may include heterogeneous noninflated and inflated cases of PTA member-states whose adoption of maximum PTA-flexibility provisions is determined by distinct theoretical processes. The noninflated observations in *Escape Flexibility*'s top-category are specifically PTA member-states that negotiate and adopt each provision step-by-step—that is, *gradationally*—from the minimal provision of SGs to eventually all four flexibility provisions listed earlier.

As such, these noninflated observations are based on the theoretically assumed ordered continuum as scholars theorize that PTA-member countries sequentially negotiate and adopt each flexibility provision step-by-step when they pursue these provisions as a *safety valve* (i.e., temporarily raise trade barriers) to insure against trade shocks engendered by trade liberalization (Rosendorff and Milner 2001; Pelc 2009). Further, these "insurance-seeking" PTA member-states—i.e., the noninflated cases—pursue the flexibility provisions step-by-step to genuinely signal to other countries that they will revert to implementing PTA-mandated trade reforms once the shocks dissipate and refrain from using escape clauses to permanently raise trade barriers (Prusa 2016, 209). By contrast, the inflated observations in *Escape Flexibility's* top-category are PTA member-states that negotiate and incorporate all four flexibility provisions for *immediate* trade protection of domestic firms. These inflated observations are not based on the theoretically assumed ordered continuum for *Escape Flexibility*, as PTA member-state governments that seek immediate trade protection for domestic firms via maximum flexibility provisions do so not for insurance but to appease domestic "special interests" who pressure these governments to *not* implement trade reforms (Bhagwati et al. 1999).

Accordingly, the inflated top-category observations are "NAs" as they are not based on *Escape Flexibility's* ordered continuum, in which the top-category is theoretically assumed to result from PTA-member countries that adopt the four flexibility provisions step-by-step for trade insurance. Further, we posit that the inflated cases *concurrently* pursue all four PTA-flexibility provisions for immediate protection. Doing so allows these governments to assure these firms that they are diligently pursuing the firms' protectionist interests, which helps them to raise campaign finance.

[1] *Regime Dummy* = 1 for PTAs where the *least* democratic member state's Polity IV score is 6 or greater.

[2] *Democratization* dummy = 1 for PTA member-states that have democratized over the past ten years.

[3] See the online appendix, table A7 for more details.

Obtaining maximum escape clauses for immediate protection concurrently is thus a secondary process distinct from the sequential step-by-step adoption of flexibility provisions theoretically assumed by Baccini et al. in their ordered *Escape Flexibility* scale.

A detailed analysis of *Escape Flexibility*'s top-category presented and illustrated in the online appendix, Section A2.1 (see online appendix, figure A1) reveals that a sizable 35 percent of PTA member-states are inflated cases who not only maintain high trade barriers but also pursued maximum flexibility provisions in PTAs concurrently for immediate protection of domestic firms. However, the remaining 65 percent (noninflated cases) negotiate and sequentially adopt flexibility provisions step-by-step for insurance against trade shocks. The online appendix, table A1, lists some examples from the data for the top-category of *Escape Flexibility* in which maximum flexibility provisions in PTAs were pursued (i) concurrently for immediate protection by some PTA member-states (inflated cases) and (ii) step-by-step by other PTA member-states (noninflated cases).

We thus analyze Baccini et al.'s data using our TiOP(C) models, which, unlike the OP model, estimate two latent equations to account for the mixture of inflated and noninflated observations in *Escape Flexibility*'s top-category. The first is the probit inflation stage, which estimates the effect of covariates on the probability that PTAs belong to the noninflated (i.e., insurance-seeking) group that pursues maximum PTA-flexibility provisions via step-by-step negotiations versus the inflated group that seeks all available flexibility provisions for the immediate protection of special interests. The second is the OP outcome-stage that estimates how another set of covariates affects the probability of each ordered category of *Escape Flexibility*, conditional on PTAs being in the noninflated group.

In the TiOP(C) models' outcome stage, we include all covariates in Baccini et al.'s main OP specification listed earlier. The TiOP(C) models' inflation-stage probit equation includes four covariates drawn from the literature on PTA-flexibility provisions: log-*GDPpc* (where *pc* = per capita), log *GDP*, *Regime Dummy*, and *Democratization*. The theoretical rationale for including these inflation-stage covariates is described in the online appendix to save space. As stated briefly here, log-*GDPpc* is included since some scholars suggest that richer PTA-member countries seek maximum PTA-escape clauses for trade protection (Prusa 2016), while others claim that they are unlikely to use maximum flexibility provisions as they can adjust to trade shocks via countercyclical fiscal policies. We add log *GDP* as economically larger member states may pursue maximum import relief measures to either "[shift] the terms of trade in their favor" (Kucik and Reinhardt 2008, 492) at the expense of smaller PTA-participating economies or to facilitate adjustment to trade shocks.

Next, *Regime Dummy*—coded as 1 for fully democratic PTA member-states—is included in the inflation stage to evaluate the following two completing claims that incumbents in PTA-member democracies pursue maximum escape flexibility measures to (i) signal to voters that they intend to shield them from terms-of-trade shocks while avoiding tariffs (Kono 2006, 374); and (ii) protect domestic industries from import competition as they are "tempted by the rents that accrue from furnishing protection" (Mansfield and Milner 2018, 374). Finally, we include the *Democratization* dummy in the TiOP(C)'s inflation stage as greater susceptibility to PTA trade reform-induced import surges may encourage recently democratized PTA-participating countries to seek maximum escape clauses in PTAs for insurance (Baccini et al. 2015, 771).

For our analyses, we first reestimate Baccini et al.'s Model 3 OP specification. We then estimate a pair of TiOP and TiOPC models (denoted TiOP1 and TiOPC1) that include (i) the aforementioned four variables in the probit inflation stage and (ii) all of Baccini et al.'s variables in the OP stage. For robustness, we add Baccini et al.'s variables to both TiOP(C) stages (denoted TiOP 2 and TiOPC 2).

The bottom half of table 1 presents the TiOP(C) models' inflation-stage re-

Table 1: Replication results for Baccini et al. (Table 3, Model 1)

	Dependent Variable: Escape Flexibility				
	OP	TiOP 1	TiOP 2	TiOPC 1	TiOPC 2
			Ordered probit stage		
Depth	0.073***	0.076***	0.071***	0.076***	0.067***
	(0.010)	(0.011)	(0.011)	(0.011)	(0.013)
Regime dummy	0.768***	0.235	0.217	0.251	0.318
	(0.143)	(0.202)	(0.200)	(0.446)	(0.299)
Depth × Regime dummy	−0.037***	−0.033***	−0.034***	−0.033***	−0.037***
	(0.011)	(0.012)	(0.012)	(0.012)	(0.014)
GDP	−0.020	0.123**	0.103*	0.119	0.156**
	(0.044)	(0.055)	(0.055)	(0.103)	(0.060)
GDPpc	0.054	−0.198*	−0.184*	−0.192	−0.301***
	(0.068)	(0.093)	(0.092)	(0.156)	(0.101)
Trade	−0.067***	−0.051*	−0.028	−0.051*	−0.012
	(0.025)	(0.030)	(0.032)	(0.030)	(0.040)
Democratization	−0.055	−0.387***	−0.405**	−0.385**	−0.411**
	(0.120)	(0.161)	(0.160)	(0.173)	(0.171)
WTO	0.036	0.222	0.312**	0.223	0.357**
	(0.113)	(0.141)	(0.147)	(0.141)	(0.163)
No. Members	−0.004	−0.002	−0.001	0.002	0.000
	(0.006)	(0.007)	(0.007)	(0.007)	(0.008)
Cut1	−0.829	0.142	−0.111	0.105	0.315
	(0.880)	(1.028)	(1.073)	(1.350)	(1.187)
Cut2	−0.289	0.745	0.488	0.708	0.874
	(0.879)	(1.027)	(1.072)	(1.353)	(1.184)
Cut3	0.221	1.374	1.109	2.068	1.486
	(0.879)	(1.028)	(1.071)	(1.378)	(1.181)
Cut4	0.741	2.105**	1.824*	2.068	2.260*
	(0.879)	(1.036)	(1.075)	(1.378)	(1.181)
			Inflation stage		
Depth			−0.008		−0.008
			(0.038)		(0.023)
Regime dummy		−1.401***	−1.509***	−1.403***	−1.435***
		(0.383)	(0.436)	(0.386)	(0.341)
Depth × Regime dummy			−0.016		−0.014
			(0.039)		(0.024)
GDP		0.438***	0.363***	0.440***	0.240**
		(0.119)	(0.134)	(0.134)	(0.101)
GDPpc		−0.753***	−0.771***	−0.761***	−0.457**
		(0.228)	(0.254)	(0.293)	(0.196)
Trade			0.119**		0.115**
			(0.060)		(0.049)
Democratization		−0.731**	−0.840**	−0.740*	−0.409
		(0.353)	(0.385)	(0.411)	(0.269)
WTO			0.477		0.330
			(0.301)		(0.211)
No. members			0.006		0.013
			(0.017)		(0.014)
Intercept		−0.932	0.568	−0.912	−0.014
		(1.341)	(2.071)	(1.423)	(1.570)
ρ				−0.032	0.684***
				(0.777)	(0.207)
Observations	559	559	559	559	559
Log-likelihood	−724.45	−704.11	−698.80	−704.11	−697.95
AIC	1,474.19	1,444.22	1,443.54	1,446.22	1,443.90
BIC	1,531.14	1,522.09	1,534.04	1,528.42	1,547.73

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

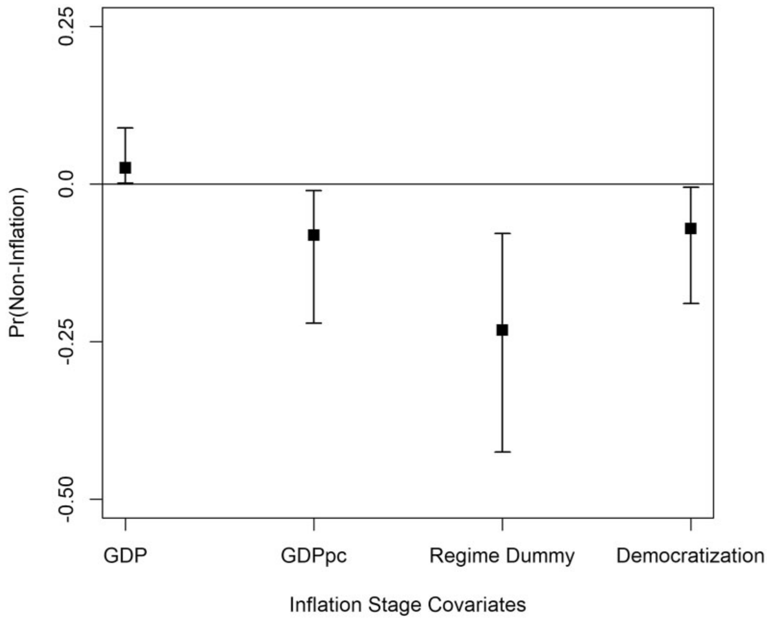

Figure 2. Effect of inflation stage covariates

sults. Using the inflation-stage estimates and parametric bootstraps,[4] we illustrate in figure 2 the effect of each covariate on the predicted probability of PTA member-states seeking maximum escape flexibility provisions for insurance (*noninflated* cases) versus immediate protection for special interests (*inflated* cases). *Regime Dummy's* negative and significant effect in the TiOP(C) inflation stage indicates that democratic PTA member-states are significantly less likely to adopt all PTA-escape flexibility measures for insurance. Instead, they pursue these provisions concurrently for the immediate protection of domestic firms. *Depth* and *Depth* × *Regime Dummy* are insignificant in the larger TiOP(C) inflation-stage specifications.

Log *GDP* is positive and significant in the TiOP(C) inflation-stage specifications, implying that economically larger PTA members pursue maximum flexibility provisions for insurance. The inflation-stage effect of *GDPpc* in figure 2 corroborates our theoretical claim that richer PTA-member countries are significantly less (conversely, more) likely to include maximum flexibility provisions for insurance (trade protection). *Democratization* reliably decreases the probability of observations being noninflated in all the TiOP(C)—barring the TiOPC 2—inflation-stage specifications. The remaining inflation-stage controls are insignificant.

Next, consider our models' ordered outcome stage results. The constituent *Regime Dummy* variable in the OP model (table 1) confirms Baccini et al.'s finding that a 0-to-1 change in *Regime Dummy* significantly increases *Escape Flexibility* when *Depth* = 0. However, once we account for *Escape Flexibility*'s top-category inflation in the TiOP(C)'s inflation stage, the *Regime Dummy* constituent term in the models' ordered outcome stages does not, unlike the OP model, have a significant effect on *Escape Flexibility*. The interaction term's two remaining components—*Depth* and *Depth* × *Regime Dummy*—remain statistically significant in the OP, TiOP, and TiOPC models' outcome stages. The effect of *Democratization* on *Escape Flexibility* in the OP model is insignificant. In the TiOP(C) models' outcome stages, however, *Democratization* has a highly significant negative effect on *Escape Flexibility*. Thus, unlike

[4] $m = 1{,}000$; all other covariates are held to means or modes.

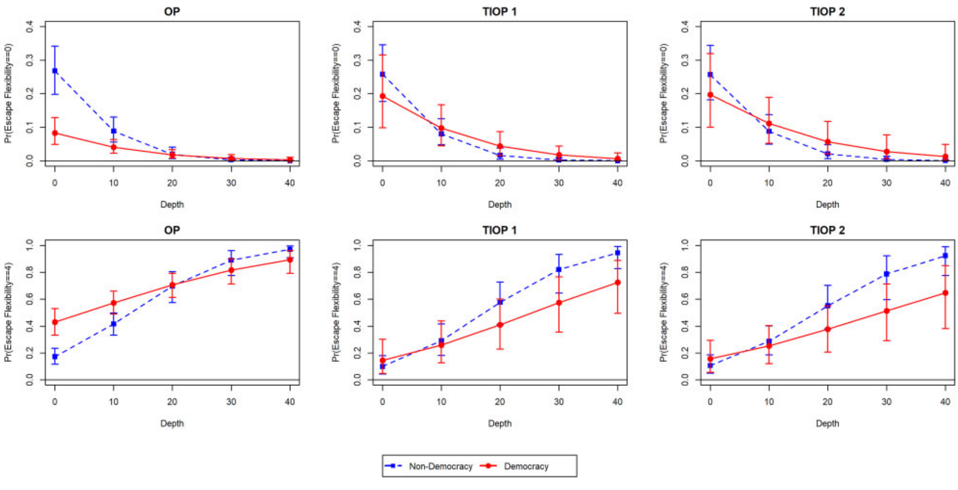

Figure 3. Effect of depth on the probability of escape flexibility by regime type

Baccini et al.'s claim, the TiOP(C) outcome stage results suggest that recently democratized PTA-member countries are *less* likely to pursue escape clauses in PTAs to hedge against uncertainty. *GDP*, *GDPpc*, and *WTO* are often significant in the TiOP(C) models' ordered outcome stages but always insignificant in the OP model.

Akaike information criterion (AIC), Bayesian information criterion (BIC),[5] and likelihood ratio (LR) test results generally favor our TiOP(C) models over Baccini et al.'s original OP model. More specifically, AICs and LR tests consistently favor all four TiOP(C) models considered over the OP model whereas BICs prefer the smaller specified TiOP(C) models over the OP model but not the larger TIOP(C) models. Given that BICs favor parsimonious models and disproportionately penalize models with larger numbers of parameters, the latter BIC result is likely attributable to our overcontrolling for *all* outcome-stage covariates in the inflation-stages of the large TiOP(C) models. AIC, BIC, and ρ-test results also generally favor the TiOP models over comparably specified TiOPC models apart from the ρ-test for our larger TiOPC specification. Our AIC and LR test results in turn favor our smaller TiOP (TiOPC) specification over our larger TiOP (TiOPC) specification. The reverse holds for our BICs, likely owing to the BIC point discussed above. In light of these results, we view the TiOP as favorable to the TiOPC but remain agnostic as to whether the larger versus smaller TiOP specification is preferable.[6]

We hence compare the marginal effects of our OP, TiOP 1, and TiOP 2 specifications by plotting the predicted probabilities of *Escape Flexibility* $= 0$ and *Escape Flexibility* $= 4$ across the range of *Depth* $\{0,10,20,30,40\}$, separately for democracies and nondemocracies.[7] The TiOP outcome stage results present each covariate's marginal effect conditional upon PTA member-states being in the noninflated set of observations in the top-category of *Escape Flexibility*. In figure 3, these estimates for *Depth* from the OP, TiOP 1, and TiOP 2 models for Pr(*Escape Flexibility*) $= 0$ and Pr(*Escape Flexibility*) $= 4$ appear along the figure's first and second rows, respectively. The plotted points from each model depict the predicted probability of seeing *Escape Flexibility* take on a particular outcome value, with 95 percent confidence intervals.

[5]For both the AIC and BIC, lower values are preferred.

[6]Given the disagreements in our AICs and BICs, as well as the small sample size and large number of parameters in this replication, we also compared our models with bias-corrected AICs (AICcs). AICcs ranked our models from most to least preferred as follows: TIOP 1, TIOP 2, TIOPC 2, TIOPC 1, and OP.

[7]We use parametric bootstraps with $m = 1{,}000$ and hold all other covariates to their means or modes.

In the OP model, greater *Depth* decreases Pr(*Escape Flexibility*)= 0 more substantially for nondemocracies than for democracies (see Baccini et al. 2015). The OP results further imply that increasing *Depth* exerts a substantial positive effect on the likelihood that a PTA includes all *Escape Flexibility* mechanisms (i.e., Pr(*Escape Flexibility*)= 4) for nondemocracies compared to democracies. This supports Baccini et al.'s prediction that the effect of *Depth* on *Escape Flexibility* is weaker for democracies compared to nondemocracies.

However, these effects dissipate in the TiOP models. When Pr(*Escape Flexibility*) = 0, figure 3 shows that *Depth's* negative effect on the likelihood that a PTA includes no flexibility mechanisms is the same for nondemocracies and democracies. Moreover, the TiOP models suggest that the effect of a 0-to-10 increase in *Depth* is indistinguishable between democracies and nondemocracies, with noticeably wider confidence intervals on each estimated effect for subsequent levels of *Depth*. Importantly, the former (0-to-10) range encompasses 72 percent of all observations for *Depth* in our data and was the range with the clearest gap in effect across Baccini et al. and our OP specifications. Hence, contrary to Baccini et al., figure 3 suggests that after statistically accounting for *Escape Flexibility's* top-category inflation via the TiOP(C) models, political regime type no longer exerts a significant moderating effect on the relationship between *Depth* and *Escape Flexibility*.

Our second application considers Girod et al.'s (2016) cross-national evaluation of repression dynamics, the resource curse, and anti-government protests. Their main theoretical prediction is that oil-rich autocracies often resort to extreme repression targeted against anti-government campaign-specific activities. Girod et al. test this prediction by employing an ordered dependent variable *Repression Intensity* from the NAVCO 2.0 dataset that ranges from none (0) to extreme repression (3). The top-category (3) operationalizes "extreme repression" against anti-government campaign-specific activities executed via "mass violence," "torture," or "kill(ing)" (Chenoweth and Lewis 2013, 13). The said top-category contains an excessive share (68 percent) of observations in *Repression Intensity*.

We argue and demonstrate in the online appendix, Section A3.1 that these excessive observations incorporate heterogeneous inflated and noninflated cases of state-perpetrated repression generated from distinct theoretical processes. Specifically, the noninflated cases in *Repression Intensity*'s top-category are observations of substantial state-driven repression *targeted* against specific anti-government opposition groups or activists, which is costly since targeting requires fine-grained information to execute (Kalyvas 2006). These noninflated cases are based on Girod et al.'s theoretically determined ordered continuum that emphasizes heavy-handed repression by autocrats against anti-government campaign-specific activities. By contrast, the inflated cases in *Repression Intensity*'s top-category are observations of extreme state-sponsored repression of noncampaign-specific activities that are cheaper and *indiscriminately* applied against a wider swath of the population, including bystanders, to typically demonstrate power (Kalyvas 2006, 146–8; Lyall 2009). As such, the inflated share of observations in *Repression Intensity*'s top-category is "NAs," as they are *not* based on Girod et al.'s theoretically determined ordered continuum, which focuses on significant targeted repression against anti-government campaign-specific activities.

As described in the online appendix, we employ information on opposition campaigns included in the Girod et al. data to empirically assess the share of noninflated and inflated cases in *Repression Intensity*'s top-category. Our analysis reveals that about 37 percent of the said measure's top-category is inflated, as these are cases of extreme yet indiscriminate repression of noncampaign-specific activities that are not based on the theorized ordered continuum of severe repression of campaign-specific activities (see online appendix, figure A2). The other 63 percent of *Repression Intensity*'s top-category observations are noninflated and based on the ordered continuum of repression of campaign-specific activities. The online appendix, table

A3, provides examples of (non)inflated cases drawn from the extreme repression top-category of *Repression Intensity* in the Girod et al. data. Given the aforementioned mixture of top-category cases, we present a replication of Girod et al.'s analysis with the OP, TiOP, and TiOPC models in the online appendix. Therein, we highlight evidence for top-category inflation and find that failure to address top-category inflation can, in some cases, *strengthen* key findings as opposed to *diminishing* them.

Conclusion

IR scholars often use OP or OL models to assess ordered outcomes such as state-perpetrated repression, exchange rate regime choice, PTA-flexibility provisions, and physical integrity rights. Yet, there are often excessive observations in the top-category of such ordinal dependent variables arising from distinct theoretical processes that are not accounted for by the OP and OL models. Building on extant methodological IR research (Xiang 2010; Bagozzi et al. 2014; Bagozzi et al. 2015), our TiOP(C) models address the preponderant share of top-category observations in ordered dependent variables. Results from Monte Carlos and two IR applications reveal that the TiOP(C) estimates are preferable to those of the OP model when an ordered dependent variable is "top-category inflated."

To this end, we have elucidated the problem of top-category inflation, highlighting the measurement and theoretical drivers of this phenomenon in multiple IR contexts. The TiOP(C) models provide a solution to the aforementioned problem by jointly estimating two latent equations that account for mixtures of inflated and noninflated cases in the top-category of ordered dependent variables. This feature is crucial because, unlike the standard OP model, the TiOP(C) estimators enable researchers to obtain more accurate estimates of covariates, detect nonmonotonic effects, and avoid model misspecification. Moreover, the fact that the TiOP(C) estimators can explicitly evaluate competing theoretical processes associated with heterogeneous top-category cases in ordered measures suggests that more theoretical work must be done to fully account for the behavior of these heterogeneous units in IR outcomes evaluated by inflated top-category ordered dependent variables. This will lead to substantively rich theories that have significant explanatory power to identify and explain why, when, and which political actors adopt tactics that determine top-category outcomes of ordered IR variables such as state-perpetrated atrocities, interstate war outcomes, and perceptions of economic threat.

Supplementary Information

Supplementary information is available in the *ISAFPA* data archive.

References

Baccini, L., A. Dür, and M. Elsig. 2015. "The Politics of Trade Agreement Design: Revisiting the Depth–Flexibility Nexus." *International Studies Quarterly* 59: 765–75.

Bagozzi, B.E, T. Brawner, B. Mukherjee, and V. Yadav. 2014. "Regional International Organizations and Individual Immigration Attitudes: Results from Finite Mixture Models.." *International Interactions* 40: 350–375.

Bagozzi, B.E., D.W. Hill, W.H. Moore, and B. Mukherjee. 2015. "Modeling Two Types of Peace: The Zero-Inflated Ordered Probit (ZiOP) Model in Conflict Research." *Journal of Conflict Resolution* 59: 728–52.

Bhagwati, J.N., P. Krishna, and A. Panagariya. 1999. *Trading Blocs: Alternative Approaches to Analyzing Preferential Trade Agreements.* Cambridge, MA: MIT Press.

Brysk, A., and A. Mehta. 2014. "Do Rights at Home Boost Rights Abroad? Sexual Equality and Humanitarian Foreign Policy." *Journal of Peace Research* 51: 97–110.

Chenoweth, E., and O.A. Lewis. 2013. Appendix Accompanying `Unpacking nonviolent campaigns: Introducing the NAVCO 2.0 Dataset. Accessed on 2/10/2022 from:https://www.du.edu/korbel/sie/media/documents/faculty_pubs/chenowethlewisjprappendix.pdf.

Girod, D.M., M.A. Stewart, and M.R. Walters. 2016. "Mass Protests and the Resource Curse: The Politics of Demobilization in Rentier Autocracies." *Conflict Management and Peace Science* 35: 503–22.
Harris, M.N., and X. Zhao. 2007. "A Zero-Inflated Ordered Probit Model, with An Application to Modelling Tobacco Consumption." *Journal of Econometrics* 141: 1073–99.
Ju, C. 2023. "Determinants of Conflict-Related Sexual Violence: A Meta-Reanalysis Distinguishing Two Classes of Zero Observations." *International Studies Quarterly* 67: sqad044.
Kalyvas, S.N. 2006. *The Logic of Violence in Civil War*. Cambridge: Cambridge University Press.
Kono, D.Y. 2006. "Optimal Obfuscation: Democracy and Trade Policy Transparency." *American Political Science Review* 100: 369–84.
Kucik, J., and E. Reinhardt. 2008. "Does Flexibility Promote Cooperation? An Application to the Global Trade Regime." *International Organization* 62: 477–505.
Lebovic, J.H. 2021. "Predilection or Prediction? Country Selection for the President's Daily Intelligence Brief, 1961–1977." *Foreign Policy Analysis* 18: orab036.
Lyall, J. 2009. "Does Indiscriminate Violence Incite Insurgent Attacks? Evidence from Chechnya." *Journal of Conflict Resolution* 53: 331–62.
Mansfield, E.D., and H.V. Milner. 2018. "The Domestic Politics of Preferential Trade Agreements in Hard Times." *World Trade Review* 17: 371–403.
Murdie, A. 2014. "The Ties That Bind: A Network Analysis of Human Rights International Nongovernmental Organizations." *British Journal of Political Science* 44: 1–27.
Pelc, K.J. 2009. "Seeking Escape: The Use of Escape Clauses in International Trade Agreements." *International Studies Quarterly* 53: 349–68.
Piazza, J.A., and J.I. Walsh. 2009. "Transnational Terror and Human Rights." *International Studies Quarterly* 53: 125–48.
Prusa, T.J. 2016. "Antidumping Provisions in Preferential Trade Agreements." In *The World Trade System: Trends and Challenges*. Cambridge, MA: MIT Presspp. 117–48.
Rosendorff, B.P., and H.V. Milner. 2001. "The Optimal Design of International Trade Institutions: Uncertainty and Escape." *International Organization* 55: 829–57.
Savun, B., and B.J. Phillips. 2009. "Democracy, Foreign Policy, and Terrorism." *Journal of Conflict Resolution* 53: 878–904.
Xiang, J. 2010. "Relevance as a Latent Variable in Dyadic Analysis of Conflict." *The Journal of Politics* 72: 484–98.

Bagozzi, Benjamin E. et al. (2024) Top-Category Inflation in Ordered International Relations Outcomes. *Foreign Policy Analysis*, https://doi.org/10.1093/fpa/orae006