

Keller, Susanne; Krautheim, Sebastian

Working Paper

Trade Negotiations Under Fire: Social Identity and the Rising Opposition to Free Trade

CESifo Working Paper, No. 12337

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Keller, Susanne; Krautheim, Sebastian (2025) : Trade Negotiations Under Fire: Social Identity and the Rising Opposition to Free Trade, CESifo Working Paper, No. 12337, Munich Society for the Promotion of Economic Research - CESifo GmbH, Munich

This Version is available at:

<https://hdl.handle.net/10419/336038>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Working Papers

Trade Negotiations under Fire: Social Identity and the Rising Opposition to Free Trade

Susanne Keller, Sebastian Krautheim

Imprint:

CESifo Working Papers

ISSN 2364-1428 (digital)

Publisher and distributor: Munich Society for the Promotion
of Economic Research - CESifo GmbH

Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740

Email office@cesifo.de
<https://www.cesifo.org>

Editor: Clemens Fuest

An electronic version of the paper may be downloaded free of charge

- from the CESifo website: www.ifo.de/en/cesifo/publications/cesifo-working-papers
- from the SSRN website: www.ssrn.com/index.cfm/en/cesifo/
- from the RePEc website: <https://ideas.repec.org/s/ces/ceswps.html>

Trade Negotiations under Fire: Social Identity and the Rising Opposition to Free Trade*

Susanne Keller[†] Sebastian Krautheim[‡]

December 2025

Abstract

Massive protests against – and spectacular failures of – international trade negotiations are an important element of the “globalization backlash” developed economies have experienced over the last two decades. We analyze a model of endogenous social identity where the government chooses between accepting or rejecting a trade agreement involving the recognition of a lower precautionary standard. We show that protests by an NGO against the trade agreement can alter the social identity equilibrium, feeding back into the political process and possibly resulting in a polarization of society. We focus on two possible outcomes: the trade agreement may either be maintained, but at lower welfare gains than initially expected (the “CETA-case”), or it may be abolished altogether (the “TTIP-case”).

Keywords: trade negotiations, trade agreements, standards, social identity, NGOs, globalization backlash, TTIP, CETA

JEL Classification: F13, F68, L31

* We would like to thank Carsten Eckel, Hartmut Egger, Johann Graf Lambsdorff, Claudius Löhnert, Gianmarco Ottaviano, Davide Sala, Thierry Verdier, Gerald Willmann and participants of 23th International Economics Workshop, Göttingen; the 34th BGPE Workshop, Augsburg; the 16th Australasian Trade Workshop, Cairns; the Workshop ‘Shaping Globalization’, JGU Mainz; the CESifo Area Conference on Global Economy 2025; the ETSG 26th Annual Conference, Milan; the 4th ECON-Workshop, University of Passau; as well as seminar participants at the University of Passau and University College Dublin for fruitful discussions and comments. Both authors gratefully acknowledge financial support by the German Research Foundation (DFG, project no. 552944111).

[†] University of Passau

[‡] Corresponding author, University of Passau and CESifo, email: sebastian.krautheim@uni-passau.de

1 Introduction

Massive protests against – and spectacular failures of – international trade negotiations are an important element of the “globalization backlash” of recent decades. The post-war period was characterized by decades of steady tariff reductions combined with the general public’s mild sympathy towards trade liberalization. More recently, however, part of the public opinion has forcefully turned against trade liberalization. Cases in point, which both motivate and guide our analysis, are the massive protests against the U.S.–EU Transatlantic Trade and Investment Partnership (TTIP) and the Canada–EU Comprehensive Economic and Trade Agreement (CETA) between 2014 and 2016, resulting in the failure of the former.¹

There is a recent and growing literature on other dimensions of the globalization backlash (see Colantone et al., 2022 for a survey). Despite the major impact they can have on trade policy, protests by NGOs and other civil society actors have received relatively little attention by economic scholars. The aim of this paper is to contribute to closing this gap.

Our starting point is the notion that a fundamental novelty in modern trade negotiations such as TTIP and CETA is that they include the recognition of (precautionary) standards. Pascal Lamy, former European Commissioner of Trade (1999-2004) and former Director-General of the WTO (2005-2013) writes: *“In the old world, [opening trade] was done by eliminating protective tariffs [...]. In the new world, this is done by harmonizing precaution.”* (Lamy, 2016, p. 4). We highlight in section 2 that concerns about the effects on regulation and standards indeed loomed large in the skepticism towards TTIP and CETA in the general public. Trade negotiations in the ‘new world’ therefore involve new issues and new stakeholders. To cite Pascal Lamy once more: *“[...] tariffs were roughly ideologically neutral. [...] Precaution is ideologically different depending on cultures, history or religion. [...] think about animal welfare, GMOs, or data privacy. [...] this has to do with [...] values. They are areas where the spectrum of preferences among different groups of people is extremely wide.”* (Lamy, 2015, p. 5). In line with these observations, we argue that ‘new world’ trade agreements touch at a fundamental question of social identity, namely “who do we want to be?” – as a group within society or as society at large.

In this paper, we address the following main questions. How do ‘new world’ trade negotiations interact with the narrative(s) of “who we want to be” within a society? Can this explain why trade agreements – even if they are widely agreed to be economically beneficial – are ultimately rejected by the general public? Why and how do NGOs play a central role in this process? When do the very same governments that negotiated and promoted the accords turn against them at the last minute – and what are the costs in terms social cohesion if they do not?

We see *social identity theory* as the natural framework to address these questions. The theory originated in social psychology (see Tajfel, 1981, Tajfel and Turner, 1979) and was

¹ We provide details on the negotiations, the massive protests as well as the about-face of key European politicians in section 2, where we also highlight that the negotiations failed *before* the 2016 U.S. presidential elections.

popularized in economics by Akerlof and Kranton (2000, 2010). Based on the concept of the social identity equilibrium introduced by Shayo (2009), Grossman and Helpman (2021) introduce social identity theory into the field of International Economics and argue that it is key for understanding modern-day trade policy. While their analysis is focused on the ‘old world’ question of setting tariffs, we use social identity theory to better understand the recognition of foreign (precautionary) standards. In our view, social identity theory also provides the ideal framework to rationalize the role of NGOs in the protests: we interpret – to our knowledge for the first time in the Economics literature – NGOs as providers of *focal points* of social identity, telling potential supporters who they *should* want to be. NGOs thereby affect the social identity equilibrium and therefore also policy choices as well as real economic outcomes.²

We analyze a model of endogenous social identity where a political process along the lines of Lindbeck and Weibull (1987), Dixit and Londregan (1996), and Grossman and Helpman (1996, 2001) results in a welfare maximizing government choosing between accepting or rejecting a trade agreement that involves the recognition of a low precautionary standard for the production of a specific good. The good can either be produced with a high or a low standard, where the foreign country has the absolute advantage in low-standard production. We consider a setting with two types of consumers. ‘Non-caring’ consumers are indifferent concerning the standard, while ‘caring’ consumers, in contrast, do care about the standard to a point that they get reduced utility from consuming a low-standard good.

Besides direct material utility from consumption, consumers derive utility from identifying with particular groups. Specifically, they can identify with their own group (caring or non-caring), with ‘society at large’ which includes both caring and non-caring consumers, and with an NGO that promotes ‘ethical’ (high-standard) production. For utility from identification with their own group and society at large, we follow Shayo (2009) and Grossman and Helpman (2021) and assume that when consumers identify with a given group, they derive utility from the level of material utility of an average member of this group but incur a dissonance cost which is the higher, the more their individual material utility differs from that of the average member of the group. While this dissonance cost is zero when identifying with the own group (caring or non-caring), it matters for identifying with society at large: a more unequal society is more costly to identify with – for both groups. Note that different from the idea of membership of a group, identifying (or not) with a group only affects the individual self-perception and does not change the size or composition of the group: society, for example, always consists of all individuals living in a country, no matter how many of them identify with it.

In addition to identifying with the own group and society at large, consumers can also choose to identify with the NGO. In our model, the NGO provides a “focal point” of social identity, which we define as a set of opinions and choices (economic, ethical and

² There is an interesting tension here to Grossman and Helpman (2021), who argue that traditional theories of the political economy of trade policy with Special Interest Groups (SIGs) lobbying for protection (e.g. Grossman and Helpman, 1994) seem less relevant for explaining modern-day trade policy. We argue, however, that when it comes to ‘new world’ trade agreements, a specific type of SIG has become much more, not less, relevant: NGOs, their campaigns and their privileged position to affect the social identity equilibrium.

political) that characterize the prototypical (“ideal”) member of the group. For some of these dimensions of the focal point, a supporter may be closer or more distant to the prototypical supporter, varying the utility she gets from identifying with the NGO. We model this with two groups of caring consumers, those fully and those partially congruent with the NGO’s demands. Other dimensions of the prototypical supporter, however, may be conditions *sine qua non*. We assume, for example, that, for its supporters, the NGO rules out identification with the group of non-caring consumers, as these are diametrically opposed to the values of the NGO represented by the prototypical supporter.³

We keep both the modeling of the NGO and the negotiation process very stylized. We seek to focus on how NGO protests triggered by trade negotiations on precautionary standards can alter the social identity equilibrium, lead to polarization and possibly make governments abandon the agreement. We therefore abstract from incentives, trade-offs and interactions both within a single NGO and between NGOs within the NGO sector.⁴

We take what Lamy (2016) calls the ‘old world’ as our starting point. We model this as a situation where previous trade negotiations have reduced tariffs to zero, but standards had not been adjusted from their autarky levels. We consider a high initial standard in the home country and a low one in the foreign country. We show that an NGO emerging into this equilibrium in the home country leaves the ‘old world’ equilibrium unaffected, except that all caring consumers start identifying with the NGO leading to additional welfare gains for them.

We then turn to the situation where the government can sign a ‘new world’ trade agreement with Foreign, implying a recognition of the foreign low standard. This would imply that Foreign, with its absolute advantage in low-standard production, would become the only producer of the product. Just as in autarky, the government takes the impact of its actions on the social identity equilibrium and the resulting effects on aggregate welfare into account. We seek to model the fact that governments were taken aback from the massive NGO protests against TTIP and CETA. We therefore assume that the government does not anticipate that the NGO might adjust its prototype, i.e., we treat the NGO response to the trade agreement as an ‘unknown unknown’ for the government.⁵

When the government announces the intention to sign the ‘new world’ trade agreement,

³ Hogg (2001), Hogg and van Knippenberg (2003) and Van Knippenberg et al. (2004) construct social identity models where prototypes in groups emerge endogenously and leaders emerge that are congruent to the group prototype. In our case, the NGO is free to define the prototype and the group of supporters forms around it.

⁴ We highlight in the conclusions how the setting could be extended to a model with many NGOs reflecting the large number of NGOs involved in the TTIP/CETA-protests. This would shift the focus of the analysis to the internal workings of the NGO sector, which seems a worthwhile endeavor in its own right.

⁵ There are different ways to think about this assumption. One is that the NGO emergence into the ‘old world’ equilibrium is a novelty and initially there is limited information on NGO technology and its impact both on the side of the government and the NGO. Another perspective is based on the observation that a large coalition of very diverse NGOs joined the TTIP and CETA protests. This includes NGOs with core issues like hunger relief, environmental protection, and food safety as well as trade unions, anti-globalization groups, Christian groups, and others. Many of the NGOs involved have core issues other than precautionary standards. Moreover, most of them had been quite indifferent to trade negotiations in the past. We therefore think that it is plausible to treat the NGO response as an ‘unknown unknown’ of the government. Clearly, one could also consider a model where the government miscalculates the level of resistance. We think of our model as the zero-one version of this approach.

from the perspective of the NGO, this changes the character of society: it is transformed from a society with production and consumption under the high standard to a society that has betrayed the NGO’s values. The NGO therefore starts a campaign and adjusts its prototype making identification with the NGO and society at large incompatible.

In the ‘old world’ equilibrium all caring consumers identify with the NGO and society at large. With the ‘new world’ trade agreement, they have to choose between the two, leading to a polarization of society. This creates losses in identity utility which the government had not accounted for. We show that when the protests break loose, two possible cases arise. In the ‘CETA-case’ aggregate welfare is lower than anticipated but still justifies ratification even if this implies a more polarized society. In the ‘TTIP-case’, welfare losses due to polarization of society outweigh the gains from the trade agreement so that the government experiences ‘negotiator’s remorse’ and turns against the agreement it had just negotiated.

This paper contributes to the literature, pioneered by Akerlof and Kranton (2000), that introduces social identity theory in the analysis of economic questions (e.g. Akerlof and Kranton, 2002, 2005, 2010; Benjamin et al., 2010; Bénabou and Tirole, 2011; Bordalo et al., 2016). Most related to our work are two recent papers that apply social identity to trade policy. As outlined in the introduction, Grossman and Helpman (2021) provide a rationale for populist trade policies using social identity in a political economy model with unilateral tariff setting. Abramson and Shayo (2022) investigate the link between international economic integration and social identity. They show that integration does not necessarily lead to a common identity and that a national identity of a very large (small) country even fosters (harms) economic integration. Other related studies use social identity to analyze preferences for income redistribution (Shayo, 2009; Klor and Shayo, 2010); the relation between tax policy and social cohesion (Harms and Niedringhaus, 2024); the rise of populism and nationalism (Besley and Persson, 2021; Bonomi et al., 2021) and generally how politicians can affect voters’ beliefs and identities (Ash et al., 2021; Balcázar, 2022).

Related to the above literature is work on the political economy of populism (see Rodrik, 2018, 2021; Guriev and Papaioannou, 2022, for an overview). Blanchard and Willmann (2022) show how populism in response to macroeconomic shocks can shift trade policy towards protectionism. The populist defines the “true” people along a certain social category as the “in-group” as opposed to the “out-group” (Müller, 2016). This process resonates with the role we give the NGO in the social identity context: it determines the prototypical NGO member and thereby defines the “in-group” and “out-group”, possibly leading to a polarization of society.

A growing literature in economics and political sciences links trade exposure to support for anti-establishment or anti-globalization politics (see, e.g., Jensen et al., 2017; Colantone and Stanig, 2018a,b; Autor et al., 2020; Che et al., 2022). Some contributions directly refer to social identity theory and consider both material utility and utility from group-identification (Baccini and Weymouth, 2021; Ballard-Rosa et al., 2022; Ferrara, 2022).

Our paper also relates to the political economy literature on deep trade agreements involving regulatory cooperation (see Maggi and Ossa, 2021 for an overview). Recent studies

show how the efficiency and ratification of such trade agreements is affected by producer lobbies (Maggi and Ossa, 2020; Maggi et al., 2022; Blanga-Gubbay et al., 2025), lobbying by trade unions (Blanga-Gubbay, 2020), regulatory differences (Parenti and Vannoorenberghe, 2025), and their design (Grossman et al., 2021; Maggi and Mrázová, 2024). Our work adds to this, considering how (a change in) social identity affects the acceptance of trade agreements involving the recognition of foreign standards.

Finally, this paper also contributes to the literature on NGOs in International Economics (Aldashev and Verdier, 2009; Aldashev et al., 2015; Krautheim and Verdier, 2016; Hatte and Koenig, 2018; Koenig et al., 2025). This literature mainly builds on the theory of *private politics* à la Baron (2001, 2003) – an Industrial Organization approach to analyze how social activists affect firm choices directly, instead of lobbying for regulation. Our paper, in contrast, analyses how NGOs affect the ratification of trade agreements and therefore the *public politics* of economic globalization.

2 Protests Against TTIP and CETA in Europe

Our analysis is motivated and guided by the massive protests against the TTIP and CETA agreements around 2015 and their impact on government choices. TTIP was on its way to become the deepest bilateral trade agreement in the history of trade negotiations and was seen as a model for future trade deals outside the multilateral approach of the World Trade Organization (WTO).⁶ After the content of the secret negotiation process was revealed, massive protests organized by broad coalitions of civil society actors erupted in many EU member states.⁷ The protests brought together a broad group of diverse NGOs and civil society actors, including unions, environmental groups, consumer protection associations, church organizations, anti-globalization and anti-capitalist initiatives, as well as other grassroots movements.⁸ The protests culminated in a series of mass demonstrations in several major German cities including Berlin with over 150,000 participants in October 2015 (The Economist, 2015). On the EU level, more than 500 European NGOs supported the “Stop TTIP” initiative, which collected over three million signatures for a European Citizenship Initiative against TTIP (Bauer, 2016, p.70; De Ville and Siles-Brügge, 2016). Along with the protests came a shift in the public perception of TTIP. While in February 2014, 55% of Germans viewed TTIP as beneficial for Germany, in February 2016, this share declined markedly to just 17% (see PEW Research Center, 2014; Bluth, 2016, p. 15).

A central criticism of TTIP and CETA concerned the harmonization or mutual recognition of regulatory standards. NGOs warned that the agreements could allow the import of products that do not meet EU norms, including hormone-treated beef, genetically modified organisms (GMOs), chlorine-washed poultry, and pesticide-treated crops (see BEUC The

⁶ Maggi and Ossa (2021) write that TTIP together with CETA and TPP (the Transpacific Trade Partnership) “are often viewed as the archetypes of modern trade agreements” (p.23).

⁷ As the anti-TTIP movement gained traction, NGOs began to also link their campaigns to CETA which had initially attracted less public attention, but contained similar provisions.

⁸ Some of the more prominent NGOs include national branches of Greenpeace, WWF, Oxfam, and foodwatch; Verdi (one of the major German trade unions) and the Seattle to Brussels Network.

European Consumer Organization, 2014, p.5-8; BUND, 2025). These practices, common in the U.S. and Canada, clashed with Europe's precautionary approach to regulation and stirred fears of weakened consumer protection and animal welfare standards (see BUND, 2015; Friends of the Earth Europe, 2015; Institute for Agriculture and Trade Policy, 2017). Focusing on normative concerns such as consumer protection, food safety, and environmental protection, NGOs presented regulatory convergence as a threat to the European way of life, and ultimately to European identity. For instance, the infamous "chlorine chicken" became a symbol for broader fears about deregulation, U.S. dominance, and the erosion of European values (Chan and Crawford, 2017, p. 695).

Moreover, civil society actors raised environmental concerns. They criticized TTIP and CETA for impeding energy transition and European climate targets (see Neslen, 2016; Stolper, 2025). Several NGOs also saw labor and financial standards at risk (see Finance Watch, 2014; Corporate Europe Observatory, 2016). Overall, the public perception was aligned with the NGOs' standpoints: a study commissioned by the German *Bertelsmann Foundation* shows that trust in EU regulatory standards appears substantially higher than in their U.S. counterparts: 69% of Germans reported greater trust in EU food safety standards, and 66% in EU environmental standards, compared to those of the United States (Bluth, 2016, pp. 19). Moreover, the same survey shows that Germans were concerned about the effect of TTIP on consumer protection, environmental standards and workers' rights/social standards. Asked how they thought these issues would be affected by TTIP, the negative views abound. Specifically, *consumer protection*: 12% positive, 48% negative; *environmental standards*: 12% positive, 46% negative; and *workers' rights/social standards*: 10% positive, 40% negative; the remainders going to *neutral* and *don't know* (see Bluth, 2016, p. 18).

Moreover, the inclusion of the investor-state dispute settlement (ISDS) clause was strongly condemned. These provisions enable foreign investors to sue states before international, private arbitration tribunals over profit losses resulting from changes in domestic regulation and legislation. Critics argued that the ISDS mechanism posed a severe threat to national sovereignty and limited the ability of the government to impose stricter precautionary standards in the future (see Gerstetter and Meyer-Ohlendorf, 2013; Foodwatch, 2017).

Another major concern was the lack of transparency of the negotiations. Confidential negotiations had previously been standard in trade talks. With the new focus on harmonizing precautionary standards, however, this was perceived as an exclusion of civil society actors from the democratic process of opinion formation and reinforced the image of TTIP as an elite-driven project (see Dullien et al., 2015, p. 6). In response to the protests, the European Commission launched unprecedented initiatives to increase transparency and public information. It published the TTIP negotiation mandate and EU negotiating proposals online and granted Members of the European Parliament access to negotiation documents (European Commission, 2014a, European Commission, 2015, p.19). Moreover, it issued several plain-language fact sheets addressing myths about TTIP (e.g., see European Commission, 2014b). Due to pressure from the U.S., however, key documents – including joint proposals and detailed records of agreement and disagreement – remained classified (De

Ville and Siles-Brügge, 2017, p. 1500). Heldt (2020) argues that this partial transparency backfired since it revealed that full disclosure was lacking and some negotiation texts remained secret.

Heightened public pressure led key elected officials to reconsider their stance on TTIP or at least their perception of the public attitudes towards TTIP with implications on the course of negotiations. Sigmar Gabriel, the German minister of economic affairs at the time and initially one of the main proponents of the TTIP accord in European politics, turned against it and declared in August 2016 that the TTIP negotiations had “de facto failed” (The Economist, 2016) – three months *before* the 2016 US presidential elections. Ultimately, TTIP negotiations were terminated by October 2016.

Unlike TTIP, after several last-minute changes, CETA was provisionally ratified in September 2017.⁹ The initial ISDS provision was replaced by a revised framework for investor protection that explicitly affirms the right to regulate, ensuring that governments can implement environmental or consumer protection measures without facing risks from investor claims (see European Commission, 2016). Several factors help to explain the different policy outcomes of TTIP and CETA. Most importantly, criticism of CETA was generally less intense. This is particularly attributable to the fact that CETA stipulates that Canadian imports must fully comply with existing EU sanitary and phytosanitary (SPS) regulations, including those concerning genetically modified organisms (see Chapter 5 of the agreement, European Union, 2017, pp. L11/39).

3 Baseline Model

We consider two largely symmetric countries, X and F . To ease the presentation, we focus on country X and only highlight where F differs from X .

3.1 Production

Both countries produce two goods, a composite good produced with constant returns to scale, which is used as the numéraire. The production of the second good is regulated by the government, which imposes either a low (l) or high (h) production standard.¹⁰ We refer to the good as the *regulated good*. Production costs c_s with $s = l, h$ are lower under the low standard: $c_h > c_l$. Under perfect competition, the price of the good is therefore given by:

$$p = c_s. \tag{1}$$

For simplicity, we assume that both countries have identical technologies for producing the regulated good under the high standard. The only asymmetry between the two countries is that F has an absolute advantage in low-standard production, i.e. $c_l^F < c_l < c_h = c_h^F$.

⁹ CETA still awaits ratification by several national parliaments in ten EU member states (as of August 2025; see European Commission, 2025). Meanwhile, all sections falling under the exclusive competence of the EU are legally binding, including the elimination of most tariffs.

¹⁰ One can think of standards regulating consumer protection e.g. for agricultural products, environmental standards, workers’ rights or social standards.

3.2 Material Utility

Population is normalized to unity and consists of two types of consumers: consumerist “non-caring” (n) and “ethical” (e) with populations shares λ_n and λ_e , respectively. While non-caring consumers are concerned with goods prices only, ethical consumers are also concerned with production standards.

Both types of consumers obtain a material utility from consumption. For non-caring consumers, this is:

$$\nu_n = q_n^\gamma q_{0n}^{1-\gamma} \quad (2)$$

where $\gamma \in (0, 1)$, q_n is quantity consumed of the good linked to the standard and q_{0n} is the quantity consumed of the numéraire. For caring consumers, material utility is given by:

$$\nu_e = [\mathbb{I}_s + \delta(1 - \mathbb{I}_s)] q_e^\gamma q_{0e}^{1-\gamma}, \quad (3)$$

where the indicator function, $\mathbb{I}_s$, is equal to 1 if the high standard is implemented and 0 otherwise. If the good is produced with the low standard, caring consumers receive only a fraction $\delta \in (0, 1)$ of utility. The value of δ can be interpreted as a measure of the gap in terms of the ‘ethical’ quality between the high and the low standard, with a low value of δ implying a large gap. The parameter δ is therefore a demand-shifter that is well-known from the quality literature.¹¹ We extend the standard notion of quality to what could be called ‘ethical quality’ - the levels of consumer protection, animal welfare or environmental and labor standards that were respected in production.¹²

The social identity literature usually refers to group status as the source of utility from identification. In economic terms, this can be translated into income or, as in Grossman and Helpman (2021), into the material utility that can be obtained from a given income level, represented by the indirect utility function. While identical in terms of income, caring and non-caring consumers in our model differ in terms of indirect utility when production occurs under the low standard. With Cobb-Douglas preferences, consumers spend a constant fraction of their income E on each good. Substituting the optimal consumption bundle into equations (2) and (3), indirect material utility is given by:

$$\nu_n = \gamma^\gamma (1 - \gamma)^{(1-\gamma)} E c_s^{-\gamma} \quad (4)$$

$$\nu_e = [\mathbb{I}_s + \delta(1 - \mathbb{I}_s)] \gamma^\gamma (1 - \gamma)^{(1-\gamma)} E c_s^{-\gamma}. \quad (5)$$

Equations (4) and (5) show that both consumers receive the same material utility if the government imposes the high standard level. Switching to low standard-production drives a wedge between non-caring and caring consumers’ material utility.

¹¹ For instance, see Baldwin and Harrigan (2011) and references therein for the incorporation of quality into trade models.

¹² Several empirical studies provide evidence that consumers care for ethical production and have a higher willingness to pay for ‘ethical’ goods; see e.g., Loureiro and Lotade (2005), Bartling et al. (2015), and Hainmueller et al. (2015).

3.3 Utility from identification

In addition to material utility, individuals obtain psychological utility from identification with social groups (*identity utility*). This psychosocial concept is based on social identity theory pioneered by Tajfel and Turner (1979). The basic idea, formalized in Shayo (2009) and used recently in Grossman and Helpman (2021), is that individuals gain utility from identification with given groups but suffer a dissonance cost the more they differ from the prototypical (typically the average) group member. As a consequence, individuals tend to identify with groups that are not too dissimilar to themselves. In our model, there are four different groups an individual could identify with: caring consumers, non-caring consumers, society at large and an NGO. Note that identifying with a group is conceptually different from a membership. Identification with a group is an individual decision. It cannot be denied by the group and the payoff is independent of the number of other individuals identifying with the same group.

Similar in spirit to Grossman and Helpman (2021), we express total individual utility for a non-caring consumer as

$$u_n = \nu_n + A + \alpha\nu_n + \mathbb{I}_n^X \left[A^X + \alpha^X \bar{\nu}^X - \beta^X \left| \nu_n - \bar{\nu}^X \right| \right] + 0 \quad (6)$$

and caring consumers' total individual utility as

$$u_e = \nu_e + A + \alpha\nu_e + \mathbb{I}_e^X \left[A^X + \alpha^X \bar{\nu}^X - \beta^X \left| \nu_e - \bar{\nu}^X \right| \right] + \mathbb{I}_\theta^N \tilde{\theta} A^N \quad (7)$$

with utility parameters $\alpha, \alpha^X, \beta^X > 0$. As outlined above, ν_n and ν_e represent material utility of the two groups from Equations (4) and (5). $A + \alpha\nu_i$ with $i \in \{n, e\}$ is total utility from identification with the *own* consumer group. It consists of an exogenous component, $A > 0$, as well as the average material utility of the group, ν_i .¹³

The part in squared brackets represents utility from identification with society at large consisting of an exogenous constant $A^X > 0$ as well as the average material utility given by

$$\bar{\nu}^X = \lambda_n \nu_n + \lambda_e [\mathbb{I}_s + (1 - \mathbb{I}_s)\delta] \nu_n, \quad (8)$$

scaled by the exogenous constant α^X . The dissonance cost of identifying with society at large are given by the individual's distance in material utility from the society average with β^X being an exogenous scaling parameter. When the high standard is implemented ($\mathbb{I}_s = 1$), indirect utility of both groups is identical and dissonance costs are zero. Under the low standard, the two groups differ and therefore, both groups differ from the society average, implying positive dissonance costs for both groups.¹⁴ The last term in eq. (7) is utility

¹³ Note that we adopt two simplifying assumptions from Grossman and Helpman (2021). First, we assume homogeneous groups so that each individual in group $i \in \{n, e\}$ is identical to the average, which implies a positive utility from identifying with one's own group with zero dissonance. Second, we rule out cross-identification implicitly assuming that dissonance costs of identifying with the opposite group is prohibitive. To avoid notational clutter, we already account for this in equations (6) and (7). The case of identification with the NGO is more subtle and is discussed in detail below.

¹⁴ Note that it is common in social identity theory that utility of identification is independent of other

caring consumers obtain from identification with the NGO. We discuss its components and origins in the next section.

3.4 NGO - focal point of identity

To our knowledge, we are the first to introduce an NGO into a setting with formally modeled social identity preferences. We think of NGOs as providers of social identity. The NGO defines a *focal point in the identity space*, i.e., a vector of properties its prototypical NGO supporter has. When identifying with the NGO, the utility obtained depends on how close the consumer is to the prototype of the NGO. An important difference to identification with the own group or with society at large is that the point of comparison is not the (exogenous) average group member, but a reference point in the identity space defined by the NGO.

We assume that support of the high standard is the central aim of the NGO. In our model, the role of the NGO is simply to define its own prototype, this includes specifying which social groups are incompatible with NGO-identification. For example, identification with the group of caring consumers is compatible with NGO-identification (when they support the high standard), while identification with the group of non-caring consumers is not.¹⁵ As for identification with society at large, the core criterion is whether the society as a whole acts in line with the NGO's values or not. This implies that the standard chosen is key. When the high standard is implemented, all consumption and production respects the high standard and identification with society at large is in line with the NGO's values (hence the indifference of NGOs to 'old world' trade agreements). This is different under the low standard, and therefore under a 'new world' trade agreement: the regulation society has given itself contradicts the NGO's values and therefore identification with the NGO and society at large are mutually exclusive.

Formally, we define a benchmark utility of identifying with the NGO as $A^N > 0$. Moreover, we assume that apart from supporting the high standard, the prototypical NGO supporter may have other properties (e.g., ideological positions, other core issues the NGO seeks to support), which caring consumers may share to different degrees. To capture this, we assume that a fraction $1 - \eta$ of caring consumers is fully congruent with the NGO prototype, implying $\tilde{\theta} = 1$ and a fraction η is only partially congruent, implying $\tilde{\theta} = \theta < 1$. The indicator variable $\mathbb{I}_{\tilde{\theta}}^N$ takes a value of one when a caring consumer with a given $\tilde{\theta}$ identifies with the NGO. This defines all elements in the last term in eq. (7).

individual's identification choices. Different from a membership, where the sum of membership decision determines the composition of the group, the composition of society is exogenous and it is unchanged whether identifying with it inspires a sense of pride in some people or not.

¹⁵ Note that in principle also caring consumers may be unsupportive of the high standard. This would arise when the benefit of the lower price outweighs the reduced utility implied by the implementation of the low standard. This is the case e.g. for a high value of δ implying that the high and the low standard are actually quite similar in the eyes of caring consumers, or, put differently, that they do not care very much about the difference in the standard. In appendix A.1, we derive the threshold c_t^e below which caring consumers favor the low standard and show that this threshold lies outside the parameter space that is relevant for our analysis, which we define by assumptions 2, 3 and 4 in appendix A.2.

3.5 Government objective

Following Grossman and Helpman (2021), we consider a process of political competition along the lines of Lindbeck and Weibull (1987), Dixit and Londregan (1996), and Grossman and Helpman (1996, 2001). This competition game results in both parties setting the instrumental policy – here the level of the production standard – in order to maximize the utilitarian welfare function.

Aggregate welfare is obtained as the sum of individual utility equations (6) and eq. (7) weighted by the population shares. It depends on the identification regime that is summarized by (r_n^X, r_e^X, r_e^N) : r_n^X represents the fraction of non-caring consumers that identify with society at large. Equivalently, for the group of caring consumers r_e^X represents the fraction identifying with society at large. Considering NGO-identification, caring consumers are split into a group that is fully congruent (fraction $1 - \eta$ with $\tilde{\theta} = 1$) and a group that is partially congruent (fraction η with $\tilde{\theta} = \theta \in (0, 1)$) to the NGO prototype. Therefore, r_e^N representing the fraction of caring consumers that identify with the NGO can be expressed as $r_e^N = \eta r_{ep}^N + (1 - \eta) r_{ef}^N$, where r_{ep}^N represents the fraction of partially congruent consumers and r_{ef}^N represents the fraction of fully congruent consumers that identify with the NGO.¹⁶ Aggregate welfare can then be expressed as:

$$\begin{aligned} U(r_n^X, r_e^X, r_e^N) = & A + (1 + \alpha) \bar{\nu}^X \\ & + \lambda_n r_n^X \left[A^X + \alpha^X \bar{\nu}^X - \beta^X (1 - \lambda_n) (\nu_n - \nu_e) \right] \\ & + \lambda_e r_e^X \left[A^X + \alpha^X \bar{\nu}^X - \beta^X (1 - \lambda_e) (\nu_n - \nu_e) \right] \\ & + \lambda_e \eta r_{ep}^N \theta A^N + \lambda_e (1 - \eta) r_{ef}^N A^N. \end{aligned} \quad (9)$$

The first line captures material utility and utility from identification with the own consumer group. A non-caring consumer that identifies, for example, with the group of non-caring consumers does not only benefit from the lower price under the low standard but also from the fact that the lower standard also increases the average material utility of the group she identifies with.

The second and the third line represent utility from identification with society as a whole for non-caring and caring consumers, respectively. Dissonance costs for group members of group $i = n, e$ are $\beta^X (1 - \lambda_i) (\nu_n - \nu_e)$.¹⁷ Dissonance costs are zero under the high standard when material utility is identical in both groups. The low standard affects both groups differently and introduces inequality into the society, making members of both groups different from the average. On the positive side, a high level of average consumption in the society delivers a positive utility of $\alpha^X \bar{\nu}^X$ for both groups. Through this channel, a caring consumer associated with society will draw some utility from a non-caring consumer's utility of consuming a low-standard good. This does not imply that the caring consumer

¹⁶ Note that until section 5.2 the regime indicators r can only take value 1 or 0 as all members of the group make the same identification choice. This only changes in section 5.2 when the NGO protests are introduced.

¹⁷ To see this, note that $\bar{\nu}^X = \lambda_n \nu_n + \lambda_e \nu_e$, that $\lambda_n + \lambda_e = 1$, that $\nu_n \geq \nu_e$ and that the absolute value was applied.

approves of the way utility has been obtained, but just likes to identify with a society in which people have a high utility, on average. By the same token, non-caring consumers identified with society are negatively affected by the utility loss of the caring consumers under the low standard. We interpret the fraction of individuals identifying with society at large as a measure of social cohesion and its inverse as a measure of polarization of society.

Line four depicts aggregate NGO identity utility obtained by caring consumers. The first term is aggregate utility from identifying with the NGO for the fraction η of caring consumers which are not fully congruent with the NGO ($\tilde{\theta} = \theta < 1$), and the second term for those that are fully congruent ($\tilde{\theta} = 1$).

3.6 Population shares

For the sake of clarity and simplicity, we assume symmetric population shares, an assumption we maintain throughout the paper:

Assumption 1. $\lambda_n = \lambda_e$.

Equation (9) implies that with equal population shares, both consumer groups suffer the same dissonance cost for a given level of inequality. Given that both groups also have identical benefits from identification, this has the great advantage that the two groups either both identify with society or both dissociate. We formally establish this result in lemma 2 below.¹⁸ This simplification seems in order, as it reduces the number of possible identification regimes that would distract attention from the actual novelty of the model: NGO protests leading a fraction of caring consumer to dissociate from society at large. We discuss asymmetric population shares in appendix section B and show that this does not affect our main results.¹⁹

3.7 The ‘old world’ and the ‘new world’

The central aim of the analysis is to better understand the role of NGO protests for the social identity equilibrium. Our analysis requires the initial ‘old world’ equilibrium to have some properties, which translate into parameter conditions. Two properties are of central importance. First, in order to be able to capture the notion of polarization in response to NGO protests, we need to start from a situation where the high standard is implemented and both groups identify with society at large. Second, we are interested in a situation where the NGO protests after the negotiations lead a fraction of caring consumers to dissociate from society at large. We therefore need a setting where implementation of the low standard does not mechanically trigger dissociation from society.

Throughout the paper, three main assumptions assure that the equilibrium features the desired properties. We provide these assumption in appendix A.2 and their formal

¹⁸ An alternative to assuming equal population shares in the above formulation would be to model the dissonance cost using a measure of inequality on which both groups put equal weight (both groups suffer the same dissonance cost from observing the same level of inequality). And indeed: $\nu_n - \nu_e$ is a measure of inequality in society and with $\lambda_e = \lambda_n$ both groups put equal weight on it.

¹⁹ A complete analysis with unequal population shares is included in the Online Appendix.

derivations in appendix A.4.1 and A.5.3 (assumption 2), A.4.1 and A.5.3 (assumption 3), and A.5.1 (assumption 4). Subsequently, we refer to the assumptions when they become relevant and explain their origins.

4 The ‘old world’ equilibrium

As the starting point of our analysis, we seek to construct an ‘old world’ equilibrium in the spirit of Lamy (2016) who argues that a decades-long process of trade negotiations had reduced trade costs but had left national regulation and standards largely unchanged. We are neither interested in the transition from autarky to the ‘old world’, nor in the effects of the progressive lowering of trade costs on regulation over time. We therefore consider a situation where trade costs are zero but the standards are based on domestic production of the regulated good, i.e. the standard in country X is set based on the value of c_l and the standard in F is set based on c_l^F , which would also be the reference point for standard-setting in autarky. We take this as our ‘old world’ starting point of the analysis. In subsections 4.1, 4.2 and 4.3 we first derive the ‘old world’ equilibrium in the absence of an NGO and add the NGO in 4.4.

4.1 Identification choice for a given standard

We consider the following timing: the government moves first and sets the standard. Consumers observe the standard and choose identification profiles that maximize their individual utilities. We solve the model by backward induction.

The two consumer types maximize their respective utilities given by equations (6) and (7). As outlined above, we consider the ‘old world’ without the NGO in section 4.1 to 4.3 and introduce the NGO only in section 4.4. We therefore disregard the NGO-term in equations (7) and (9) for the time being.

With Cobb-Douglas preferences, consumers spend a constant fraction, γ , of their income on the regulated good and optimally choose the groups to identify with, conditional on the standard the government imposes. Both consumer types always identify with the own consumer group, as there are positive gains of $A^X + \alpha\nu_i$ at zero cost. Abstracting from the NGO in this subsection, we only have to derive conditions under which consumers choose to associate with society as a whole.

In the ‘old world’ and in the absence of the NGO, all individuals of group i choose to identify with the society in country X ($r_i^X = 1$), when utility from pride and self esteem of identification is larger or equal the dissonance costs. This is the case when in eq. (9) the terms in squared brackets are positive, i.e. when $A^X + \alpha^X \bar{\nu}^X \geq \beta^X (1 - \lambda_i) (\nu_n - \nu_e)$. When inequality within society, represented by $(\nu_n - \nu_e)$ is too strong, the condition is violated, dissonance costs outweigh the benefits and the group dissociates from society at large ($r_i^X = 0$). Note that inequality only arises under the low standard: when the government implements the high standard, caring and non-caring consumers receive the same material utility, i.e. ν_n and ν_e in equations (4) and (5) are equal and inequality is zero. This implies:

Lemma 1. *In the ‘old world’ equilibrium without an NGO, when the regulated good is produced under the high standard, there is no inequality between non-caring and caring consumers and both choose to identify with society as a whole.*

Proof. In the text.

If the low standard is imposed by the government, the cost advantage of low-standard production leads to lower prices implying a higher material utility for non-caring consumers, ν_n . While caring consumers also benefit from the lower price, they suffer from the implementation of the lower standard. This is captured by the discount factor δ in eq. (3) and introduces inequality between the two groups. This inequality makes both groups differ from the average member of society introducing a dissonance cost of identification with society at large. When these dissonance costs grow too large, individuals cease to identify with society. Given that within groups all individuals are identical, this would imply four possible identification patterns: both identify, none identifies or only one of the two identifies with society at large. Assumption 1, however, assures that only two identification patterns can arise. The reason is that for $\lambda_n = \lambda_e$, the two terms in squared brackets in eq. (9) are identical and are therefore either both positive or both negative. This allows to state the following lemma:

Lemma 2. *In the ‘old world’ equilibrium without an NGO and symmetric population shares, assumption 1 ($\lambda_n = \lambda_e$), implies that under the low standard, either all individuals identify with society as a whole (r_n^X, r_e^X) = (1, 1), or nobody does (r_n^X, r_e^X) = (0, 0).*

Proof. In the text.

Lemmas 1 and 2 allow us to move to a graphical illustration. Lemma 1 implies that for production under the high standard, both groups always identify with society, implying that for aggregate welfare (eq. (9)) the case $U^{(1,1)}(c_h)$ is relevant while the case $U^{(0,0)}(c_h)$ is not. Moreover, it follows from lemma 2 that, under the low standard, assumption 1 implies that $U^{(1,1)}(c_l)$ and $U^{(0,0)}(c_l)$ are relevant, while mixed cases (0, 1) or (1, 0) never arise (see appendix B for the case where they do).

The overall situation can be illustrated using fig. 1, which contains the main elements of our analysis. It plots the domestic production cost under the low standard c_l against the aggregate utility depending on the identification patterns as well as the level of the standard implemented.

The horizontal line represents aggregate utility $U^{(1,1)}(c_h)$ from eq. (9) (without the NGO terms in the last line) for the case that the high standard is implemented and both groups identify with society at large. Under the low standard, the cost level c_l is relevant and aggregate utilities are represented by the two convex curves $U^{(1,1)}(c_l)$ and $U^{(0,0)}(c_l)$.

Different elements shape the curves. On the one hand, lower production costs c_l lead to lower prices, which benefits both consumer groups. On the other hand, caring consumers see their utility reduced by the factor δ in eq. (3) when the low standard is implemented. Assumption 2 assures that the former effect dominates the latter for both curves leading to the decreasing profile of aggregate utilities in the graph.

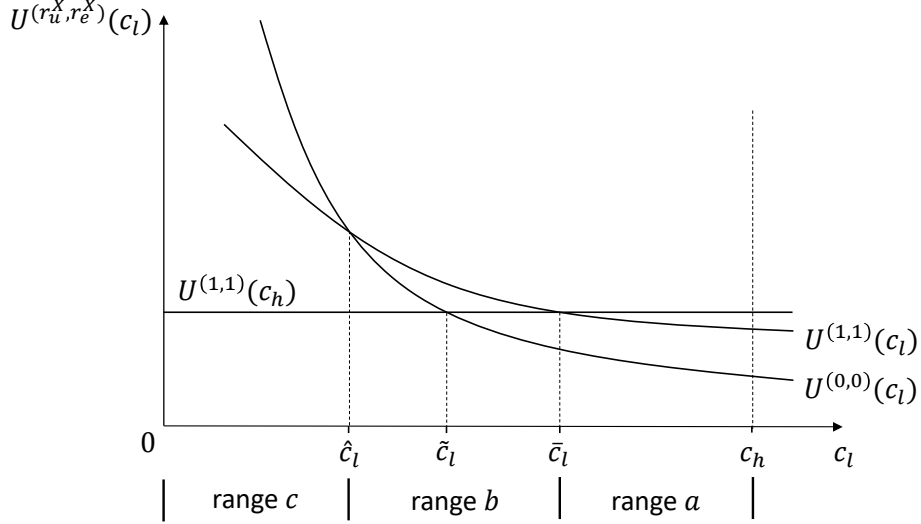


Figure 1: ‘Old world’ equilibrium without an NGO

When the low standard is implemented, lower levels of c_l increase inequality and strengthen the incentive to dissociate from society at large. We define the dissociation cutoff $\hat{c}_l$ below which inequality is so strong that consumers choose to dissociate from society at large and social cohesion breaks down. At the cutoff, we have $U^{(1,1)}(\hat{c}_l) = U^{(0,0)}(\hat{c}_l)$, implying

$$\hat{c}_l = \left(\frac{0.5\Phi [\beta^X(1-\delta) - \alpha^X(1+\delta)]}{A^X} \right)^{\frac{1}{\gamma}}, \quad (10)$$

with $\Phi = \gamma^\gamma(1-\gamma)^{1-\gamma}E$. Note that assumption 2 assures an interior solution for the cutoff (see appendix A.4.1). A higher dissociation cutoff implies that social cohesion is more fragile and breaks down even for relatively high production costs under the low standard. Anything that increases the dissonance costs increases the cutoff: a higher β^X scales the impact of inequality on the dissonance cost; a smaller δ makes the low standard more painful for caring consumers, increasing inequality and thereby the dissonance costs; and higher expenditure on the regulated good, Φ , also increases inequality and therefore dissonance under the low standard. On the positive side, the parameters A^X and α^X increase the benefits of identification with society and thereby decrease the cost level at which social cohesion breaks down.

4.2 Government decision for given identification patterns

As outlined above, the political process leads the government to maximize welfare – consisting of both material and identity utility of both consumer groups. The government moves first and takes into account how the choice of the standard affects inequality and therefore affects the choice of consumers to identify – or not – with society at large. We proceed in two steps. In this subsection, we analyze which standard maximizes welfare for given identification patterns (dissociation, or not, from society at large). In the following subsection we characterize the resulting equilibrium.

Depending on the identification pattern, there is a cutoff for c_l below which the low standard leads to higher welfare. For the case where both groups identify (1, 1), we label this cutoff $\bar{c}_l$ and for the case where both groups dissociate (0, 0) we label it $\tilde{c}_l$.

In case (1, 1), the government is indifferent between the high and low production standard when $U^{(1,1)}(c_h) = U^{(1,1)}(\bar{c}_l)$. Solving for $\bar{c}_l$ delivers the following expression:

$$\bar{c}_l = c_h \left(\frac{0.5 \left[(1 + \alpha + \alpha^X) (1 + \delta) - \beta^X (1 - \delta) \right]}{1 + \alpha + \alpha^X} \right)^{\frac{1}{\gamma}}. \quad (11)$$

Assumption 2 assures an interior solution for $\bar{c}_l$ (see appendix A.4.2). The cutoff cost level depends negatively on β^X since a stronger weight on inequality that arises under the low standard requires higher cost savings to compensate for this inequality. The cutoff increases in δ because a large δ makes production under the low standard less painful for caring consumers.

In case (0, 0), where all consumers dissociate from society, a lower standard delivers higher welfare if $c_l < \tilde{c}_l$ where the cutoff is pinned down by $U^{(1,1)}(c_h) = U^{(0,0)}(\tilde{c}_l)$. Solving for $\tilde{c}_l$ delivers:

$$\tilde{c}_l = \left[\frac{(1 + \alpha)(1 + \delta)0.5\Phi}{(1 + \alpha + \alpha^X)\Phi c_h^{-\gamma} + A^X} \right]^{\frac{1}{\gamma}}. \quad (12)$$

One last aspect of fig. 1 is the ranking of the cutoffs of $\hat{c}_l < \tilde{c}_l < \bar{c}_l$. This occurs when $U^{(1,1)}(c_h) < U^{(1,1)}(\hat{c}_l)$, which places a lower bound on c_h . We show in appendix A.4.4 that the ranking is assured by assumption 3.

4.3 ‘Old world’ equilibrium

Recall that as our starting point, we seek to construct an ‘old world’ in the spirit of Lamy (2016) who argues that a decade-long process of trade negotiations had reduced trade costs but had left national regulation and standards largely unchanged. We therefore now turn to the optimal standards and the resulting identification patterns emerging in a pre-liberalization world, i.e., a world in which the production of the regulated good is domestic. Combining the insights of the two previous subsections, the following proposition summarizes equilibrium standards and identification patterns in the ‘old world’:

Proposition 1. *There exist three cost ranges which determine the ‘old world’ equilibrium without an NGO:*

- (i) *For $\bar{c}_l < c_l$ the government implements the high standard and consumers identify with society at large (range a).*
- (ii) *For $c_l \in (\hat{c}_l, \bar{c}_l)$ the government implements the low standard and consumers identify with society at large (range b).*
- (iii) *For $c_l \in (0, \hat{c}_l)$ the government implements the low standard and consumers do not identify with society at large (range c).*

Proof. See appendix A.4.

By our symmetry assumptions, the same pattern holds for the foreign country, which features identical cutoffs. For trade negotiations involving recognition of standards, these standards need to be different in the first place. We therefore consider the case where the difference in the production cost under the low standard ($c_l^F < c_l$) implies a high standard in country X and a low standard in country F in the ‘old world’ equilibrium.²⁰

4.4 ‘Old world’ equilibrium with NGO

We now add the NGO to the ‘old world’ equilibrium. As outlined in section 2, the strong interest of NGOs in trade negotiations is a relatively recent phenomenon so that the process leading to low trade costs in the ‘old world’ was by and large unaffected by large scale NGO interventions. In this section we introduce an NGO into the ‘old world’ equilibrium of country X and show that it is essentially unaffected.

As outlined in section 3.4, the NGO offers consumers a focal point in the identity space. This means that in addition to identifying with the own group and society at large, consumers can also choose to identify with the NGO. Different from the other groups, where the reference point for inequality is the average group member, the reference point for the NGO is the prototypical group member defined by the NGO. This includes defining which social groups are incompatible with NGO-identification. As country X has already implemented the high standard, all of the regulated good consumed had been produced under the high standard in the ‘old world’ equilibrium without the NGO. When the NGO emerges into this equilibrium, identification with society at large and with the NGO are not mutually exclusive. In the absence of a cost of identification with the NGO, all caring consumers identify with it in the ‘old world’ equilibrium.

It follows from the additive structure of eq. (9) that the identification of caring consumers with the NGO leaves all other identity choices and also the optimal standard set by the government unaffected. Aggregate welfare increases as now the last line of eq. (9) is positive. From the perspective of the government, the emergence of an NGO into the ‘old world’ equilibrium is unambiguously positive.

Lemma 3. *The emergence of the pro-precautionary NGO in the context of the ‘old world’ equilibrium,*

- (i) *leads all caring consumers to identify with the NGO, increasing their utility;*
- (ii) *leaves the non-caring consumers as well as all (other) consumption and identification choices of consumers as well as the governments choice of the standard unaffected.*

Proof. In the text.

²⁰ Note that in order to keep the setting simple, we assume that for the regulated good there is a slight home bias leading consumers to choose the domestic good over the foreign good when prices are identical. This assures that production under the high standard occurs in X despite the assumption of $c_h = c_h^F$. Combined with symmetry of the two countries, this assumption assures that any cost level considered is always in the same range (a , b , or c) in both countries, avoiding additional and largely uninformative cutoffs.

In the ‘old-world’ equilibrium, the emergence of an NGO constitutes a windfall profit from the perspective of the government: nothing changes compared to the initial situation except that caring consumers obtain additional identity utility. As outlined in the introduction, we are interested in the case where the NGO sector - here represented by the NGO, responds much more strongly than the government had expected: in the TTIP and CETA protests, a large number of very diverse NGOs, many of which had previously been indifferent towards trade policy, responded with strong protests to the result of the negotiations. Here, we model this by assuming that the government expects an unchanged prototype of the NGO and the fact that it may change in response to the outcome of the negotiation is an ‘unknown unknown’ for the government.

5 Trade Agreement, NGO Protests and the Polarization of Society

We now turn to the analysis of the ‘new-world’ trade agreement. We first consider the situation where the government in country X takes the prototype of the NGO as given and considers a trade agreement with country F which includes the recognition of the foreign (low) standard. We first analyze under which conditions the government would sign the agreement. We then consider a situation where the NGO responds with protests adjusting its prototype and thereby altering the post-agreement social identity equilibrium.

5.1 The ‘new world’-trade agreement

We consider a situation where standard trade costs including tariffs are zero, so that differences in standards are the only remaining impediments to international trade. Without loss of generality, we analyze the case where in the ‘old world’ equilibrium, country X imposes the high standard and country F imposes the low standard. For simplicity, we consider a trade agreement, where the government in X simply chooses whether or not to accept the foreign standard. The idea is that the lower standard is less binding in all dimensions, so that there is no point for country F to recognize the higher standard in country X and mutual recognition is not an issue.²¹

We first analyze the case where the government takes the NGO prototype as given. In this case, the situation is fully characterized by proposition 1 combined with lemma 3, except that the relevant alternative to produce under the high standard at cost c_h is not the domestic cost of low-standard production, c_l , but the foreign cost, c_l^F , with $c_l^F < c_l$.

Figure 2 illustrates the government’s choices (FTA or no FTA) together with the expected identification patterns with society at large (identification or dissociation) as well as the relevant welfare functions for different values of c_l^F (ranges a , b and c). When the foreign production cost under the low standard is too high, specifically, when $c_l^F > \bar{c}_l$ (range a),

²¹ As outlined above, we keep the trade negotiation side deliberately simple in our model. If one wanted to place a stronger focus there, the analysis could be extended to study mutual recognition. One could, for example, introduce another sector in which country F starts off with the higher standard and then study under which conditions mutual recognition of standards is an equilibrium outcome.

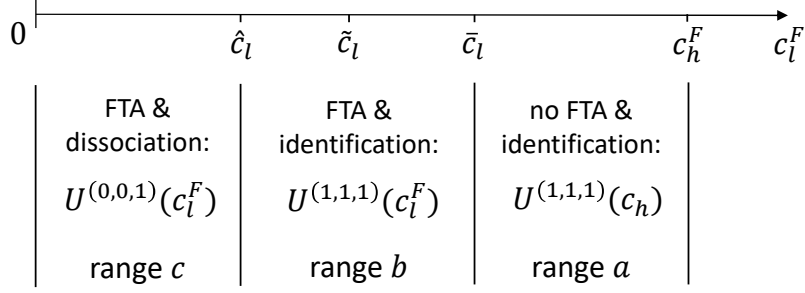


Figure 2: FTA equilibrium

the government does not sign the trade agreement and production remains in X under the high standard. When the cost advantage of low-standard production is strong enough, $c_l^F < \bar{c}_l$, the government signs the agreement recognizing the foreign standard and social cohesion – i.e. identification with society at large – is either preserved (range b) or breaks down (range c).

Ratification affects aggregate welfare through two channels. First, importing the foreign low-standard good generates standard gains from trade due to the lower foreign price. However, caring consumers obtain less utility per unit from consuming the foreign low-standard good than from consuming the high-standard good, which is governed by the discount factor δ in eq. (5). Hence, the overall effect on average material utility is positive if and only if the foreign low-standard cost advantage is sufficiently high.

Second, accepting the trade agreement benefits non-caring consumers more than caring ones. This inequality increases in the low-standard cost advantage of Foreign and we have to distinguish between the following two cases: the impact of the agreement on inequality is moderate and leaves the identification pattern unaffected, $c_l^F \in (\hat{c}_l, \bar{c}_l)$ (range b); and the impact on inequality is sufficiently strong to trigger polarization (both consumers dissociate from society), $c_l^F < \hat{c}_l$ (range c).

Range b is central for our analysis. Over the whole cost range, the government in country X favors recognizing the foreign standard, expecting social cohesion to hold. In fact, for cost levels in range b , which are above $\tilde{c}_l$, the government only favors the foreign standard *because* it expects social cohesion to hold). NGO campaigns, which we analyze in the following section, can alter social cohesion under the trade agreement and thereby aggregate welfare and possibly government choices. We therefore focus our analysis on range b .

The following proposition summarizes the main results:

Proposition 2. *Social cohesion and ratification of the trade agreement:*

- (a) *There exists a range $c_l^F \in (\hat{c}_l, \bar{c}_l)$ (range b) in which*
 - (i) *cost savings from foreign low-standard production are sufficiently high such that the ratification of the trade agreement maximizes aggregate welfare,*
 - (ii) *inequality between non-caring and caring consumers (in terms of material utility) is sufficiently small such that social cohesion of society is maintained and no polarization occurs.*

(b) *Within range b, there is a cutoff $\tilde{c}_l > \hat{c}_l$ given by eq. (12), above which the trade agreement only maximizes aggregate welfare when social cohesion is maintained. That is, in the range $c_l^F \in (\tilde{c}_l, \bar{c}_l)$ social cohesion is instrumental for the government's perspective on trade agreements.*

Proof. In the text.

5.2 NGO campaigns

We now turn to the NGO's reaction to the government's intention to recognize the lower foreign standard. From the NGO's perspective, this changes the character of society: it is transformed from a society with high-standard production and 'ethical' consumption into a society with a low standard that is not aligned with the NGO's ideals anymore. Therefore, the NGO declares that identification with the NGO and with society at large are mutually exclusive. This adjustment of the focal point takes the form of a campaign to inform and convince consumers. In the 'old world' equilibrium, all caring consumers identified with the NGO and society at large – now they have to choose between the two.

We have seen above that the introduction of the NGO into the 'old world' equilibrium left the previous identification choices and therefore also the optimal choices of the government unaffected and simply generated windfall profits in terms of identity utility. The government takes these windfall profits as given. Proposition 2 highlights that this generates a range of $c_l^F \in (\tilde{c}_l, \bar{c}_l)$, for which the trade agreement is signed only because social cohesion is expected to be maintained. The NGO forcing the caring consumers to choose between the NGO and society at large may therefore alter the fundamentals on which the initial government decision of signing the agreement was based, opening the possibility of 'negotiators regret' – a government regretting the very trade agreement it had just negotiated.

In our setting, there are three possible outcomes when the NGO declares the incompatibility with society at large: two corner solutions where all caring consumers identify with society at large and none with the NGO – or the other way round. In the main text, we focus on the interior solution where the group of caring consumers is divided between NGO and society.²² In our model, this division is generated by the parameter $\theta \in (0, 1)$, which implies that a fraction η of caring consumers obtains less utility from identifying with the NGO than the remaining fraction (see eq. (9)). As a result, caring consumers differ in terms of their trade-off between identification with society at large vs. identification with the NGO.

Figure 3 illustrates this case. The cutoff $\bar{c}_l$, which shapes the government's decision to favor the agreement, is determined by the function $U^{(1,1,1)}(c_l^F)$, i.e., it is valid when identification with society and the NGO is maintained after signing the agreement. With NGO protests, caring consumers are split into a fraction η which continues to identify with society at large (but not with the NGO) and the fraction $1 - \eta$ that chooses the opposite. Hence, all caring consumers lose some identification utility. The identification regime becomes $(r_n^X, r_e^X, r_e^N) = (1, \eta, 1 - \eta)$ and aggregate utility is given by $U^{(1,\eta,1-\eta)}(c_l^F)$.

²² We show in appendix A.6 that our main results hold also in the corner solutions.

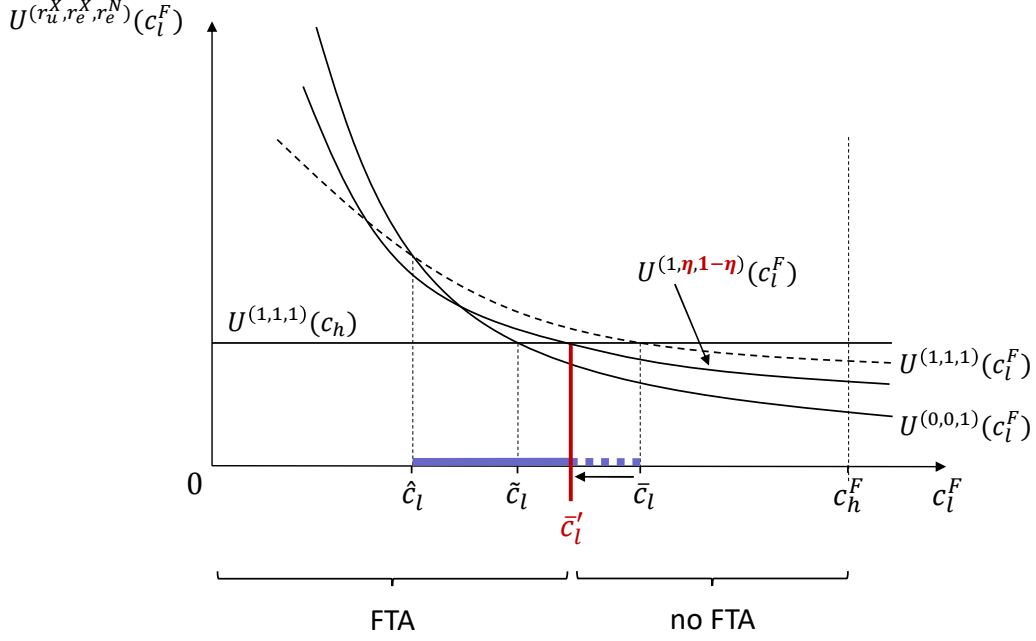


Figure 3: Effect on aggregate utility and ratification cutoff

The new ratification cutoff $\bar{c}_l'$ is determined by $U^{(1,1,1)}(c_h) = U^{(1,\eta,1-\eta)}(\bar{c}_l')$.²³ Solving for $\bar{c}_l'$ delivers:

$$\bar{c}_l' = \left[\frac{0.5\Phi \left[(1+\alpha)(1+\delta) - (1+\eta)0.5 \left(\beta^X(1-\delta) - \alpha^X(1+\delta) \right) \right]}{(1+\alpha+\alpha^X)\Phi(c_h)^{-\gamma} + (1-\eta)0.5A^X + 0.5A^N\eta\theta} \right]^{\frac{1}{\gamma}}. \quad (13)$$

This allows us to state the central result of our paper:

Proposition 3. *In the relevant cost range of $c_l^F \in (\tilde{c}_l, \bar{c}_l)$, the unanticipated NGO protests lead to a loss in identity utility for caring consumers with non-caring consumers unaffected. This has the following implications:*

- (i) *The lower aggregate utility under the trade agreement leads to a decrease in the ratification cutoff from $\bar{c}_l$ in eq. (11) to $\bar{c}_l'$ in eq. (13), with $\bar{c}_l' < \bar{c}_l$.*
- (ii) *For the range of production costs in Foreign of $c_l^F \in (\bar{c}_l', \bar{c}_l)$, an initially welfare-maximizing trade agreement is rejected (TTIP-case).*
- (iii) *For the range of production costs in Foreign of $c_l^F \in (\hat{c}_l, \bar{c}_l')$, a trade agreement is still welfare-maximizing, however, gains are lower than expected and society is polarized (CETA-case).*
- (iv) *A wider gap between the high and the low standard (lower δ) leads to a lower value of $\bar{c}_l'$ and therefore makes the TTIP-case occur for a larger range of foreign cost levels*

²³ Note that when NGO protests are successful and the government decides not to sign the agreement (and hence impose the high standard), there is no permanent damage of NGO protests on social identity utility. Most notably, joint identification with the NGO and society at large is possible and aggregate utility is again given by $U^{(1,1,1)}(c_h)$.

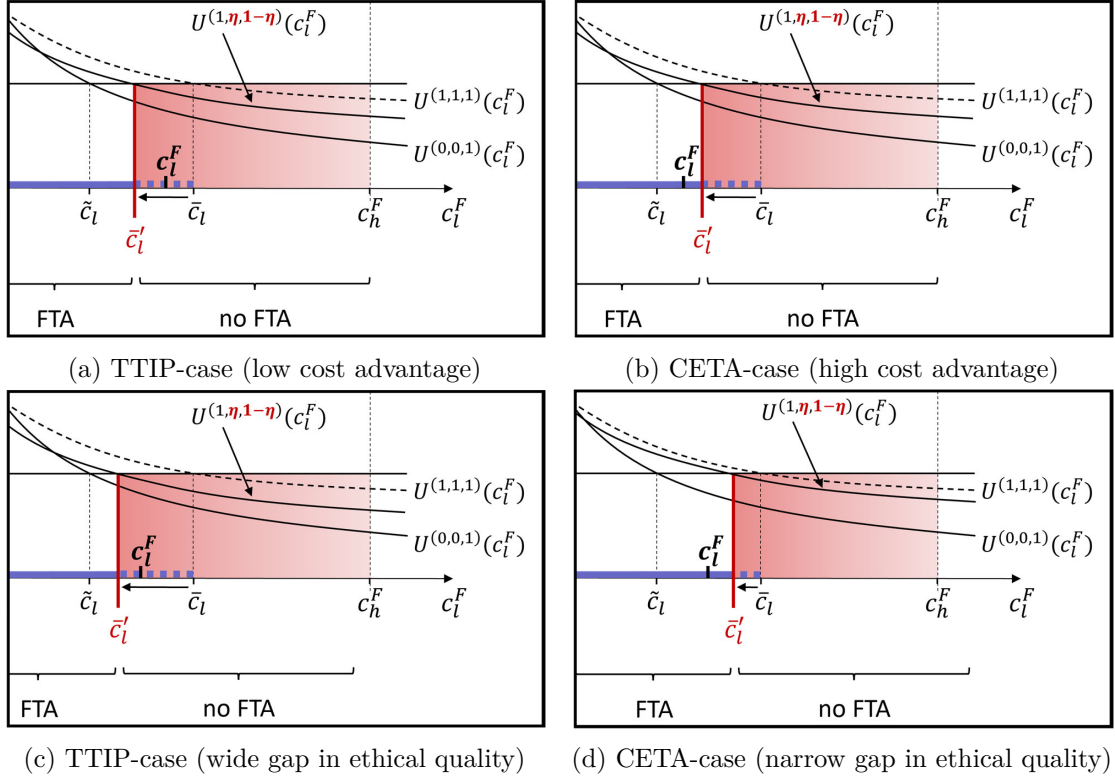


Figure 4: TTIP- and CETA-cases for low and high cost advantages of low standard production (proposition 3 (ii) and (iii)) and for wide and narrow gaps in terms of ethical quality between the high and the low production standard (proposition 3 (iv)).

$$(\partial \bar{c}_l' / \partial \delta > 0).$$

Proof. See appendix A.5.

Part (i) of the proposition is illustrated in fig. 3. The unambiguous fall in aggregate welfare leads to a downward shift of the ratification cutoff. This occurs in the cost range of $c_l^F \in (\tilde{c}_l, \bar{c}_l)$ for which the maintenance of social cohesion is central for the trade agreement to be beneficial (proposition 2b). Two cases may now occur. In fig. 4, we zoom into the core part of fig. 3. In the upper left graph we illustrate the TTIP case (point (ii) in the above proposition) where the foreign cost level c_l^F is relatively high and lies in the range between $\bar{c}_l'$ and $\bar{c}_l$ (dotted, fat purple line). In this case, the unexpected NGO protests erode social cohesion, destroying the grounds on which the government had favored the agreement. This mirrors the about-face of European (especially the German) governments when it came to ratifying the TTIP agreement between the EU and the US. Our model with social identity delivers a possible explanation: an unexpected polarization of society as a result of the NGO protests.

The graph on the upper right of fig. 4, illustrates the CETA-case (point (iii) in the above proposition). The graph is identical to the TTIP-case except that here the foreign cost level c_l^F is below $\bar{c}_l'$ (solid, fat purple line), where the cost savings of signing the agreement outweigh the welfare lost due to NGO protests. It can be seen in the graph that even though welfare exceeds the ‘old world’ level, due to the polarization of society, the advantage over

the initial situation is smaller than initially expected by the government. This mirrors the case of the CETA agreement which was barely passed after massive protests.

The two upper graphs in fig. 4 illustrate the TTIP-case and the CETA-case based on differences in the cost levels of low-standard production. Equation (13) implies that not only differences in c_l^F can tip the balance in favor of the TTIP- or the CETA-case. With the different TTIP and CETA experiences in mind, the most relevant in our context is the gap in terms of ethical quality between the high standard in country X and the low standard in Foreign, governed by the parameter δ in eq. (5). In the lower two graphs in fig. 4, we consider an identical cost level c_l^F , but different gaps between the standards. The lower left panel illustrates a wide gap in ethical quality (low value of δ), leading to a strong shift of the ratification cutoff, placing us in the TTIP-case. The intuition is that a wide gap in the standards places a heavy burden on the caring consumers when the agreement is signed. Ratification therefore requires higher cost savings, reflected in a lower ratification cutoff. The graph on the lower right features the same cost level, but with a more narrow gap in the standards (higher δ). In this case, foreign production under the low standard places a smaller burden on the caring consumers implying a smaller drop in the ratification cutoff, leading to the CETA-case.

6 Conclusion

Motivated by the massive protests against ‘new world’ trade agreements, in this paper, we analyze a model of endogenous social identity where ethical consumers care about ethical/precautionary standards respected in production of a consumption good, while consumerist consumers do not. Consumers have different identification choices to make: identification with the own group, with society at large and with an NGO promoting the high standard. As the size of the consumer groups as well as the composition of society are given, so is the average group member. That is, the reference point for identification, the prototype of the group, is constant. The NGO, however, can guide its (potential) supporters on how to respond the changes in the environment by altering the focal point of identification, i.e., its prototypical supporter.

We start from an ‘old world’ equilibrium in which tariffs are negligible, but precautionary standards differ across countries. The government can choose to accept the low foreign standard or maintain the high domestic one. Making this choice, the government takes the prototypes of all groups as given. While this is correct for the consumer groups and society at large, the NGO may adjust its prototype when the low foreign standard is recognized. In this case, from the NGO perspective, society is opposing the NGO’s values, leading the NGO to adjust its prototype such that identification with the NGO and society at large are incompatible. We argue that the case where the government does not anticipate the NGO’s reaction is the relevant one. For one thing, scale and scope of the TTIP and CETA protests exceeded the involved governments’ expectations by far. For another, this seems plausible when we switch from ‘old world’ agreements on tariffs – towards which the NGO is indifferent – to ‘new world’ agreements including the recognition of standards.

We show that protests by the NGO, adjusting its prototype and thereby forcing caring consumers to choose between society at large and the NGO, leads to a more polarized social identity equilibrium. We show that the implied losses in identity utility, associated to polarization, can lead to a situation where the government turns against the agreement (the TTIP-case) or to a ratification but with lower welfare than expected (the CETA-case).

Our model remains explicitly stylized along two dimensions: the negotiation process between the two countries and the NGO sector. In our analysis, we abstract from any negotiation process and model the government decision as a take it or leave it offer. This allows us to highlight how the recognition of a foreign standard triggering NGO protests can lead to shifts in the social identity equilibrium and alter the government's optimization problem. Future work could integrate (possibly a more stylized version of) our social identity model into a setting with multiple trade barriers, multiple standards and multiple sectors in the negotiation process. In such a context, one could also investigate how different choices concerning the negotiation process (content of the TTIP and CETA negotiations were a well-kept secret for most of the process) interact with NGO activity and possibly affect the strength of the NGO sector's reaction.

Future work could also paint a more complete picture of the NGO sector. One could introduce a large number of NGOs with explicit cost-benefit considerations for each of them and possibly externalities in their (campaign) choices. This may allow endogenizing elements like the degree of polarization when the trade agreement is signed or the scope of the protests (with negotiations on precaution triggering larger-scope protests).

References

- Abramson, B. and M. Shayo (2022). “International integration and social identity”. *Journal of International Economics* 137, p. 103577. DOI: [10.1016/j.jinteco.2022.103577](https://doi.org/10.1016/j.jinteco.2022.103577).
- Akerlof, G. A. and R. E. Kranton (2000). “Economics and Identity”. *The Quarterly Journal of Economics* 115 (3), pp. 715–753.
- Akerlof, G. A. and R. E. Kranton (2002). “Identity and Schooling: Some Lessons for the Economics of Education”. *Journal of Economic Literature* 40 (4), pp. 1167–1201. DOI: [10.1257/002205102762203585](https://doi.org/10.1257/002205102762203585).
- Akerlof, G. A. and R. E. Kranton (2005). “Identity and the Economics of Organizations”. *The Journal of Economic Perspectives* 19 (1), pp. 9–32. DOI: [10.1257/0895330053147930](https://doi.org/10.1257/0895330053147930).
- Akerlof, G. A. and R. E. Kranton (2010). *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*. Princeton University Press. DOI: [10.2307/j.ctt7rqsp](https://doi.org/10.2307/j.ctt7rqsp).
- Aldashev, G., M. Limardi, and T. Verdier (2015). “Watchdogs of the Invisible Hand: NGO monitoring and industry equilibrium”. *Journal of Development Economics* 116, pp. 28–42. DOI: [10.1016/j.jdeveco.2015.03.006](https://doi.org/10.1016/j.jdeveco.2015.03.006).
- Aldashev, G. and T. Verdier (2009). “When NGOs go global: Competition on international markets for development donations”. *Journal of International Economics* 79 (2), pp. 198–210. DOI: [10.1016/j.jinteco.2009.07.007](https://doi.org/10.1016/j.jinteco.2009.07.007).
- Ash, E., S. Mukand, and D. Rodrik (2021). *Economic Interests, Worldviews, and Identities: Theory and Evidence on Ideational Politics*. NBER Working Paper 29474. National Bureau of Economic Research. DOI: [10.3386/w29474](https://doi.org/10.3386/w29474).
- Autor, D., D. Dorn, G. Hanson, and K. Majlesi (2020). “Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure”. *American Economic Review* 110 (10), pp. 3139–83. DOI: [10.1257/aer.20170011](https://doi.org/10.1257/aer.20170011).
- Baccini, L. and S. Weymouth (2021). “Gone For Good: Deindustrialization, White Voter Backlash, and US Presidential Voting”. *American Political Science Review* 115 (2), pp. 550–567. DOI: [10.1017/S0003055421000022](https://doi.org/10.1017/S0003055421000022).
- Balcázar, C. F. (2022). “A theory of protectionist populism: The role of elite cues and identity on protectionism”. mimeo. DOI: [10.2139/ssrn.4380501](https://doi.org/10.2139/ssrn.4380501).
- Baldwin, R. and J. Harrigan (2011). “Zeros, Quality, and Space: Trade Theory and Trade Evidence”. *American Economic Journal: Microeconomics* 3 (2), pp. 60–88.
- Ballard-Rosa, C., A. Jensen, and K. Scheve (2022). “Economic Decline, Social Identity, and Authoritarian Values in the United States”. *International Studies Quarterly* 66 (1). DOI: [10.1093/isq/sqab027](https://doi.org/10.1093/isq/sqab027).
- Baron, D. P. (2003). “Private Politics”. *Journal of Economics & Management Strategy* 12 (1), pp. 31–66. DOI: [10.1111/j.1430-9134.2003.00031.x](https://doi.org/10.1111/j.1430-9134.2003.00031.x).

- Baron, D. P. (2001). “Private Politics, Corporate Social Responsibility, and Integrated Strategy”. *Journal of Economics & Management Strategy* 10 (1), pp. 7–45. DOI: [10.1111/j.1430-9134.2001.00007.x](https://doi.org/10.1111/j.1430-9134.2001.00007.x).
- Bartling, B., R. A. Weber, and L. Yao (2015). “Do Markets Erode Social Responsibility?” *The Quarterly Journal of Economics* 130 (1), pp. 219–266. DOI: [10.1093/qje/qju031](https://doi.org/10.1093/qje/qju031).
- Bauer, M. (2016). *Manufacturing discontent: the rise to power of anti-TTIP groups*. ECIPE Occasional Paper 02/2016. URL: <https://ecipe.org/wp-content/uploads/2016/11/Manufacturing-Discontent.pdf>.
- Bénabou, R. and J. Tirole (2011). “Identity, Morals, and Taboos: Beliefs as Assets”. *The Quarterly Journal of Economics* 126 (2), pp. 805–855. DOI: [10.1093/qje/qjr002](https://doi.org/10.1093/qje/qjr002).
- Benjamin, D. J., J. J. Choi, and A. J. Strickland (2010). “Social Identity and Preferences”. *The American Economic Review* 100 (4), pp. 1913–1928. DOI: [10.1257/aer.100.4.1913](https://doi.org/10.1257/aer.100.4.1913).
- Besley, T. and T. Persson (2021). “The Rise of Identity Politics: Poilcy, Political Organization, and Nationalist Dynamics”. mimeo.
- BEUC The European Consumer Organization (2014). *Food and the Transatlantic Trade & Investment Partnership (TTIP)*. URL: https://www.beuc.eu/publications/beuc-x-2014-030_ipa_beuc_position_paper_ttip_food.pdf (visited on 06/25/2025).
- Blanchard, E. and G. Willmann (2022). “Unequal gains, prolonged pain: A model of protectionist overshooting and escalation”. *Journal of International Economics* 135, p. 103559. DOI: [10.1016/j.jinteco.2021.103559](https://doi.org/10.1016/j.jinteco.2021.103559).
- Blanga-Gubbay, M. (2020). “Trade Unions and Trade Protection”. mimeo.
- Blanga-Gubbay, M., P. Conconi, and M. Parenti (2025). “Lobbying for Globalisation”. *The Economic Journal* 135 (666), pp. 487–518. DOI: [10.1093/ej/ueae082](https://doi.org/10.1093/ej/ueae082).
- Bluth, C. (2016). *Attitudes to global trade and TTIP in Germany and the United States*. Bertelsmann Stiftung.
- Bonomi, G., N. Gennaioli, and G. Tabellini (2021). “Identity, Beliefs, and Political Conflict”. *The Quarterly Journal of Economics* 136 (4), pp. 2371–2411. DOI: [10.1093/qje/qjab034](https://doi.org/10.1093/qje/qjab034).
- Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016). “Stereotypes”. *The Quarterly Journal of Economics* 131 (4), pp. 1753–1794. DOI: [10.1093/qje/qjw029](https://doi.org/10.1093/qje/qjw029).
- BUND (2025). *TTIP – Freiheit für die Gentechnik*. URL: https://www.bund.net/fileadmin/user_upload_bund/publikationen/ttip_und_ceta/ttip_und_ceta_freiheit_fuer_gentechnik.pdf (visited on 06/25/2025).
- BUND (2015). *TTIP und die Chemikalienpolitik in der EU. Mögliche Auswirkungen auf bestehende und künftige Regulierungen*. URL: https://www.bund.net/fileadmin/user_upload_bund/publikationen/ttip_und_ceta/ttip_und_ceta_chemiepolitik_eu.pdf (visited on 06/25/2025).
- Chan, A. T. and B. K. Crawford (2017). “The puzzle of public opposition to TTIP in Germany”. *Business and Politics* 19 (4), pp. 683–708. DOI: [10.1017/bap.2017.32](https://doi.org/10.1017/bap.2017.32).

- Che, Y., Y. Lu, J. R. Pierce, P. K. Schott, and Z. Tao (2022). “Did trade liberalization with China influence US elections?” *Journal of International Economics* 139, p. 103652. DOI: [10.1016/j.jinteco.2022.103652](https://doi.org/10.1016/j.jinteco.2022.103652).
- Colantone, I., G. Ottaviano, and P. Stanig (2022). “The backlash of globalization”. *Handbook of International Economics: International Trade*. Ed. by G. Gopinath, E. Helpman, and K. Rogoff. Vol. 5. Elsevier, pp. 405–477. DOI: [10.1016/bs.hesint.2022.02.007](https://doi.org/10.1016/bs.hesint.2022.02.007).
- Colantone, I. and P. Stanig (2018a). “Global Competition and Brexit”. *American Political Science Review* 112 (2), pp. 201–218. DOI: [10.1017/S0003055417000685](https://doi.org/10.1017/S0003055417000685).
- Colantone, I. and P. Stanig (2018b). “The Trade Origins of Economic Nationalism: Import Competition and Voting Behavior in Western Europe”. *American Journal of Political Science* 62 (4), pp. 936–953. DOI: [10.1111/ajps.12358](https://doi.org/10.1111/ajps.12358).
- Corporate Europe Observatory (2016). *The Great CETA Swindle*. URL: <https://corporateeurope.org/sites/default/files/attachments/great-ceta-swindle.pdf> (visited on 06/25/2025).
- De Ville, F. and G. Siles-Brügge (2016). *TTIP: Wie das Handelsabkommen den Welthandel verändert und die Politik entmacht*. Bielefeld: transcript Verlag. DOI: [10.1515/9783839434123](https://doi.org/10.1515/9783839434123).
- De Ville, F. and G. Siles-Brügge (2017). “Why TTIP is a game-changer and its critics have a point”. *Journal of European Public Policy* 24, pp. 1491–1505. DOI: [10.1080/13501763.2016.1254273](https://doi.org/10.1080/13501763.2016.1254273).
- Dixit, A. and J. Londregan (1996). “The Determinants of Success of Special Interests in Redistributive Politics”. *Journal of Politics* 58 (4), pp. 1132–1155. DOI: [10.2307/2960152](https://doi.org/10.2307/2960152).
- Dullien, S., A. Garcia, and J. Janning (2015). *A Fresh Start for TTIP*. 124. European Council on Foreign Relations, pp. 1–11.
- European Commission (2016). *CETA: EU and Canada agree on new approach on investment in trade agreement*. Press release. URL: https://ec.europa.eu/commission/presscorner/detail/en/ip_16_399 (visited on 07/29/2025).
- European Commission (2025). *EU-Canada Comprehensive Economic and Trade Agreement (CETA)*. URL: https://policy.trade.ec.europa.eu/eu-trade-relationships-country-and-region/countries-and-regions/canada/eu-canada-agreement_en (visited on 06/24/2025).
- European Commission (2014a). *Statement: TTIP: ‘I’m delighted that EU governments decided to make the TTIP negotiating mandate public’ – says De Gucht*. STATEMENT/14/306. URL: https://ec.europa.eu/commission/presscorner/detail/en/STATEMENT_14_306 (visited on 06/26/2025).
- European Commission (2014b). *The top 10 myths about TTIP – Separating fact from fiction*. Publications Office of the European Union. DOI: [10.2781/185592](https://doi.org/10.2781/185592).

- European Commission (2015). *Trade for All: Towards a More Responsible Trade and Investment Policy*. European Commission, 2015. Luxembourg: Publications Office of the European Union.
- European Union (2017). *Comprehensive Economic and Trade Agreement (CETA) between Canada and the European Union and its Member States*. URL: https://eur-lex.europa.eu/eli/agree_international/2017/37/oj (visited on 07/07/2025).
- Ferrara, F. M. (2022). “Why does import competition favor republicans? Localized trade shocks and cultural backlash in the US”. *Review of International Political Economy*, pp. 1–24. DOI: [10.1080/09692290.2021.1980898](https://doi.org/10.1080/09692290.2021.1980898).
- Finance Watch (2014). *Financial Services in TTIP? Understanding Finance #2*. URL: <https://www.finance-watch.org/wp-content/uploads/2018/10/understanding-finance-ttip.pdf> (visited on 06/25/2025).
- Foodwatch (2017). *The Impact of CETA on the Environment, Climate and Health*. URL: <https://www.foodwatch.org/en/campaigns/free-trade-agreements/the-impact-of-ceta-on-the-environment-climate-and-health> (visited on 07/11/2025).
- Friends of the Earth Europe (2015). *How TTIP Undermines Food Safety and Animal Welfare*. URL: https://friendsoftheearth.eu/wp-content/uploads/2015/02/briefing-ttip-food-safety-feb2015_1.pdf (visited on 06/25/2025).
- Gerstetter, C. and N. Meyer-Ohlendorf (2013). *Investor-state dispute settlement under TTIP – a risk for environmental regulation?* Heinrich-Böll-Stiftung and Ecologic Institute, pp. 1–25.
- Grossman, G. M. and E. Helpman (1996). “Electoral Competition and Special Interest Politics”. *The Review of Economic Studies* 63 (2), pp. 265–286. DOI: [10.2307/2297852](https://doi.org/10.2307/2297852).
- Grossman, G. M. and E. Helpman (2021). “Identity Politics and Trade Policy”. *The Review of Economic Studies* 88 (3), pp. 1101–1126. DOI: [10.1093/restud/rdaa031](https://doi.org/10.1093/restud/rdaa031).
- Grossman, G. M. and E. Helpman (1994). “Protection for Sale”. *American Economic Review* 84 (4), pp. 833–850.
- Grossman, G. M. and E. Helpman (2001). *Special Interest Politics*. Cambridge, Massachusetts: MIT Press.
- Grossman, G. M., P. McCalman, and R. W. Staiger (2021). “The “New” Economics of Trade Agreements: From Trade Liberalization to Regulatory Convergence?” *Econometrica* 89 (1), pp. 215–249. DOI: [10.3982/ECTA17536](https://doi.org/10.3982/ECTA17536).
- Guriev, S. and E. Papaioannou (2022). “The Political Economy of Populism”. *Journal of Economic Literature* 60 (3), pp. 753–832. DOI: [10.1257/jel.20201595](https://doi.org/10.1257/jel.20201595).
- Hainmueller, J., M. J. Hiscox, and S. Sequeira (2015). “Consumer Demand for Fair Trade: Evidence from a Multistore Field Experiment”. *Review of Economics and Statistics* 97 (2), pp. 242–256. DOI: [10.1162/REST_a_00467](https://doi.org/10.1162/REST_a_00467).

- Harms, P. and J. Niedringhaus (Jan. 2024). *On Social Cohesion and Social Disintegration*. Working Papers 2401. Gutenberg School of Management and Economics, Johannes Gutenberg-Universität Mainz.
- Hatte, S. and P. Koenig (2018). “The Geography of NGO Activism against Multinational Corporations”. *The World Bank Economic Review* 34 (1), pp. 143–163. DOI: [10.1093/wber/lhy007](https://doi.org/10.1093/wber/lhy007).
- Heldt, E. C. (2020). “Contested EU trade governance: Transparency conundrums in TTIP negotiations”. *Comparative European Politics* 18 (2), pp. 215–232. DOI: [10.1057/s41295-019-00183-4](https://doi.org/10.1057/s41295-019-00183-4).
- Institute for Agriculture and Trade Policy (2017). *CETA: European Food and Agriculture Standards Under Threat*. URL: https://www.iatp.org/sites/default/files/2017-09/0-CETA_Summary_sept17_06w.pdf (visited on 06/26/2025).
- Jensen, J. B., D. P. Quinn, and S. Weymouth (2017). “Winners and Losers in International Trade: The Effects on US Presidential Voting”. *International Organization* 71 (3), pp. 423–457. DOI: [10.1017/S0020818317000194](https://doi.org/10.1017/S0020818317000194).
- Klor, E. F. and M. Shayo (2010). “Social identity and preferences over redistribution”. *Journal of Public Economics* 94 (3), pp. 269–278. DOI: [10.1016/j.jpubeco.2009.12.003](https://doi.org/10.1016/j.jpubeco.2009.12.003).
- Koenig, P., S. Krautheim, C. Löhnert, and T. Verdier (2025). “Local global watchdogs: Trade, sourcing and the internationalization of social activism”. *Journal of International Economics* 157, p. 104116. DOI: [10.1016/j.jinteco.2025.104116](https://doi.org/10.1016/j.jinteco.2025.104116).
- Krautheim, S. and T. Verdier (2016). “Offshoring with endogenous NGO activism”. *Journal of International Economics* 101, pp. 22–41. DOI: [10.1016/j.jinteco.2016.03.009](https://doi.org/10.1016/j.jinteco.2016.03.009).
- Lamy, P. (2016). *The Changing Landscape of International Trade*. The Frank D. Graham Lecture. URL: <https://pascallamy.eu.files.wordpress.com/2017/02/2016-04-07-lamy-princeton-graham-lecture-final.pdf>.
- Lamy, P. (2015). *The New World of Trade*. Jan Tumlrir Lecture. URL: <https://pascallamy.eu.wordpress.com/2015/03/09/colloque-le-bien-commun-futur-paradigme-de-la-gouvernance-de-mers/>.
- Lindbeck, A. and J. W. Weibull (1987). “Balanced-Budget Redistribution as the Outcome of Political Competition”. *Public Choice* 52 (3), pp. 273–297.
- Loureiro, M. L. and J. Lotade (2005). “Do fair trade and eco-labels in coffee wake up the consumer conscience?” *Ecological Economics* 53 (1), pp. 129–138. DOI: [10.1016/j.ecolecon.2004.11.002](https://doi.org/10.1016/j.ecolecon.2004.11.002).
- Maggi, G. and M. Mrázová (2024). *Harmonization... What Else? The Role for International Regulatory Agreements*. NBER Working Paper 33318. National Bureau of Economic Research, Inc.

- Maggi, G., M. Mrázová, and J. P. Neary (Feb. 2022). “Choked By Red Tape? The Political Economy Of Wasteful Trade Barriers”. *International Economic Review* 63 (1), pp. 161–188. DOI: [10.1111/iere.12533](https://doi.org/10.1111/iere.12533).
- Maggi, G. and R. Ossa (2020). *Are Trade Agreements Good For You?* NBER Working Papers 27252. National Bureau of Economic Research, Inc.
- Maggi, G. and R. Ossa (2021). “The Political Economy of Deep Integration”. *Annual Review of Economics* 13 (1), pp. 19–38. DOI: [10.1146/annurev-economics-121020-032425](https://doi.org/10.1146/annurev-economics-121020-032425).
- Müller, J.-W. (2016). *What Is Populism?* Philadelphia: University of Pennsylvania Press. DOI: [10.9783/9780812293784](https://doi.org/10.9783/9780812293784).
- Neslen, A. (2016). *Leaked TTIP Energy Proposal Could ‘Sabotage’ EU Climate Policy*. URL: <https://www.theguardian.com/environment/2016/jul/11/leaked-ttip-energy-proposal-could-sabotage-eu-climate-policy> (visited on 06/25/2025).
- Parenti, M. and G. Vannoorenberghe (2025). “A Simple Theory of Deep Trade Integration”. *Journal of the European Economic Association* 23 (3), pp. 891–940. DOI: [10.1093/jeea/jvae046](https://doi.org/10.1093/jeea/jvae046).
- PEW Research Center (2014). *Support in Principle for US-EU Trade Pact – But Some Americans and Germans Wary of TTIP*. Details. Washington, D.C.: PEW Research Center. URL: <https://www.pewresearch.org/global/2014/04/09/support-in-principle-for-u-s-eu-trade-pact/> (visited on 07/21/2025).
- Rodrik, D. (2018). “Populism and the economics of globalization”. *Journal of International Business Policy* 1 (1), pp. 12–33. DOI: [10.1057/s42214-018-0001-4](https://doi.org/10.1057/s42214-018-0001-4).
- Rodrik, D. (2021). “Why Does Globalization Fuel Populism? Economics, Culture, and the Rise of Right-Wing Populism”. *NBER Working Paper Series*.
- Shayo, M. (2009). “A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution”. *American Political Science Review* 103 (2), pp. 147–174. DOI: [10.1017/S0003055409090194](https://doi.org/10.1017/S0003055409090194).
- Stolper, E.-C. (2025). *CETA und Klimaschutz*. URL: <https://www.bund.net/themen/ttip-ceta/ceta-und-klimaschutz/> (visited on 06/25/2025).
- Tajfel, H. and J. C. Turner (1979). “The social psychology of intergroup relations”. *An integrative theory of intergroup conflict*. Ed. by W. G. Austin and S. Worchel.
- The Economist (2015). *The TTIP of the spear: Selling Europe’s trade agreement with America as “strategic” has problems*. The Economist. URL: <https://www.economist.com/europe/2015/10/17/the-ttip-of-the-spear> (visited on 07/05/2025).
- The Economist (2016). *Why Germans are protesting against free trade. A nation of exporters is increasingly wary of liberalising trade*. URL: <https://www.economist.com/the-economist-explains/2016/09/15/why-germans-are-protesting-against-free-trade> (visited on 05/14/2025).

A Appendix

A.1 Caring consumers' preference for the high standard

If low-standard production yields sufficiently high cost savings, caring consumers prefer the low standard. We therefore derive a threshold c_l^e for the low-standard production cost above which caring consumers strictly prefer the high standard. The threshold c_l^e derives from equalizing material utility under the low and the high standard, that is $\nu_e(c_l^e) = \nu_e(c_h) \leftrightarrow c_l^e = \delta^{\frac{1}{\gamma}} c_h$. This implies caring consumer prefer the high standard iff:

$$c_l > c_l^e = \delta^{\frac{1}{\gamma}} c_h \quad (\text{A.1})$$

Therefore, our analysis is based on the cost range $c_l > c_l^e$ to assure that we are in the relevant parameter space.

A.2 Parameter restrictions

Here we collect the assumptions which assure that we are in the relevant parameter space for our analysis. Their formal derivations are provided in appendix A.4.1 and A.5.3 (assumption 2), A.4.1 and A.5.3 (assumption 3), and A.5.1 (assumption 4).

Assumption 2. $\frac{\alpha^X(1+\delta)}{1-\delta} < \beta^X < \frac{(1+\alpha+\alpha^X)(1-\delta)+\alpha^X 0.5\eta\theta(1+\delta)}{(1-\delta)(1+0.5\eta\theta)}$.

Note that for $\delta = \frac{1+\alpha}{1+\alpha+2\alpha^X}$ the left-hand side and right-hand side of the inequality are identical. The inequality holds for smaller values of δ . Assumption 2 therefore implies a sufficiently high perceived gap in ethical quality between the high and the low standard, i.e. $\delta < \frac{1+\alpha}{1+\alpha+2\alpha^X}$.

Assumption 3.

$$\left(\frac{(1+\alpha+\alpha^X)\Phi[\beta^X(1-\delta)-\alpha^X(1+\delta)]}{A^X[(1+\alpha+\alpha^X)(1+\delta)-\beta^X(1-\delta)]+A^N\eta\theta[\beta^X(1-\delta)-\alpha^X(1+\delta)]} \right)^{\frac{1}{\gamma}} < c_h < \left(\frac{0.5\Phi(\beta^X(1-\delta)-\alpha^X(1+\delta))}{\delta A^X} \right)^{\frac{1}{\gamma}}$$

with $\Phi = \gamma^\gamma(1-\gamma)^{(1-\gamma)}E$.

Assumption 4. $A^N < A^X$

A.3 Properties of aggregate utility

Note that eq. (9) can be rewritten as:

$$\begin{aligned} U(r_n^X, r_e^X, r_e^N) &= A + (1+\alpha)\bar{\nu}^X \\ &\quad + \lambda_n \mathbb{I}_n^X \left[A^X + \alpha^X \bar{\nu}^X - \beta^X(1 - [\mathbb{I}_s + (1 - \mathbb{I}_s)\delta])(1 - \lambda_n)\nu_n \right] \\ &\quad + \lambda_e \mathbb{I}_e^X \left[A^X + \alpha^X \bar{\nu}^X - \beta^X(1 - [\mathbb{I}_s + (1 - \mathbb{I}_s)\delta])(1 - \lambda_e)\nu_n \right] \\ &\quad + \lambda_e \left(\eta\theta A^N + (1 - \eta)A^N \right) \end{aligned} \quad (\text{A.2})$$

Differentiating aggregate utility (eq. (A.2)) in the ‘old world’ and before NGO emergence with respect to c_s delivers:

$$\begin{aligned} \frac{\partial U(r_n^X, r_e^X)}{\partial c_s} = & -\gamma 0.5 \Phi c_s^{-\gamma-1} \{ (1 + \alpha) (1 + [\mathbb{I}_s + (1 - \mathbb{I}_s)\delta]) \\ & + \mathbb{I}^X [\alpha^X (1 + [\mathbb{I}_s + (1 - \mathbb{I}_s)\delta]) - \beta^X (1 - [\mathbb{I}_s + (1 - \mathbb{I}_s)\delta])] \} \end{aligned} \quad (\text{A.3})$$

which is negative iff

$$(1 + \alpha) (1 + [\mathbb{I}_s + (1 - \mathbb{I}_s)\delta]) + \mathbb{I}^X [\alpha^X (1 + [\mathbb{I}_s + (1 - \mathbb{I}_s)\delta]) - \beta^X (1 - [\mathbb{I}_s + (1 - \mathbb{I}_s)\delta])] > 0 .$$

According to lemmas 1 and 2, we have to consider the following cases: the regulated good is produced under the high standard and all consumers identify with society at large ($s = h$ and $(1, 1)$), the regulated good is produced under the low standard and all consumers identify with society ($s = l$ and $(1, 1)$), and the regulated good is produced under the low standard and consumers do not identify with society ($s = l$ and $(0, 0)$) which implies that $\mathbb{I}_s = 0$ and $\mathbb{I}^X = 0$.

It is obvious to see that the derivative is negative for c_h and $(1, 1)$ and c_l and $(0, 0)$ since aggregate utility does not include dissonance costs in these two cases. For the case c_l and $(1, 1)$, the derivative is negative iff

$$\frac{(1 + \alpha + \alpha^X)(1 + \delta)}{(1 - \delta)} > \beta^X . \quad (\text{A.4})$$

It follows that aggregate utility $U(r_n^X, r_e^X)$ decreases strictly monotonically in the cost level c_s for all possible combinations of production and identification patterns if eq. (A.4) holds. Which is always the case under assumption 2.

A.4 ‘Old world’ cutoff cost levels

A.4.1 Dissociation cutoff

Consumers are indifferent to identification with society at large if the benefits from identification (i.e. the group status) exactly offset disutility in terms of inequality between non-caring and caring consumers. Note that this approach is equivalent to setting $U^{(1,1)}(c_l) = U^{(0,0)}(\hat{c}_l)$ which reduces to the starting point of the derivation of $\hat{c}_l$ above. Formally, this is

$$A^X + \alpha^X \bar{\nu}^X(\hat{c}_l) = \beta^X (1 - \delta) 0.5 \nu_n(\hat{c}_l) .$$

The cutoff level is therefore given by:

$$\hat{c}_l = \left(\frac{0.5 \Phi [\beta^X (1 - \delta) - \alpha^X (1 + \delta)]}{A^X} \right)^{\frac{1}{\gamma}} . \quad (\text{A.5})$$

For $c_l < \hat{c}_l$ consumers dissociate from society. By definition $c_l > 0$. Hence, for $\hat{c}_l$ being an interior cutoff we need $\beta^X > \frac{\alpha^X(1+\delta)}{1-\delta}$, which constitutes the first element of assumption 2. For uniqueness, we derive utility from identification with society with respect to c_l , which strictly monotonically increases in the cost level c_l iff $\frac{\alpha^X(1+\delta)}{1-\delta} < \beta^X$. Therefore, assumption 2 assures existence and uniqueness of the dissociation cutoff $\hat{c}_l$.

Moreover, we need to ensure that $\hat{c}_l$ lies above c_l^e . This can be easily done by plugging eq. (A.5) into eq. (A.1) which delivers the following condition. There exists a cost range $c_l^e \leq \hat{c}_l < c_l$, in which consumers identify with society at large and caring consumers always prefer the high standard, under the following condition:

$$c_h < \left(\frac{0.5\Phi \left(\beta^X(1-\delta) - \alpha^X(1+\delta) \right)}{\delta A^X} \right)^{\frac{1}{\gamma}}. \quad (\text{A.6})$$

This condition constitutes the second element of assumption 3.

A.4.2 Cutoff for the low standard: case (1,1)

For the case (1,1), where both groups identify with society, the government implements the low standard when $U^{(1,1)}(c_h) < U^{(1,1)}(c_l)$.

Derivation of the cutoff. Setting $U^{(1,1)}(c_h) = U^{(1,1)}(c_l)$ and solving for c_l delivers the cutoff $\bar{c}_l$:

$$\bar{c}_l = c_h \left(\frac{0.5 \left[(1 + \alpha + \alpha^X)(1 + \delta) - \beta^X(1 - \delta) \right]}{1 + \alpha + \alpha^X} \right)^{\frac{1}{\gamma}}. \quad (\text{A.7})$$

For $c_l < \bar{c}_l$ the low standard maximizes aggregate welfare.

Existence of an interior cutoff. By definition, $c_l \in (0, c_h)$. To prove that $\bar{c}_l \in (0, c_h)$, we consider the limits of c_l . Therefore, we solve eq. (A.7) for c_h and compare $U^{(1,1)}(c_h)$ to $U^{(1,1)}(c_l)$ with $c_l = \{0, c_h\}$ for a given value of c_h to find the optimal standard when c_l converges to its limits.

$$\lim_{c_l \rightarrow 0} c_l \left(\frac{1 + \alpha + \alpha^X}{0.5 \left[(1 + \alpha + \alpha^X)(1 + \delta) - \beta^X(1 - \delta) \right]} \right)^{\frac{1}{\gamma}} = 0.$$

As $c_h > 0$, the low standard is preferred when c_l approaches zero.

$$\lim_{c_l \rightarrow c_h} c_l \left(\frac{1 + \alpha + \alpha^X}{0.5 \left[(1 + \alpha + \alpha^X)(1 + \delta) - \beta^X(1 - \delta) \right]} \right)^{\frac{1}{\gamma}} = c_h \left(\frac{1 + \alpha + \alpha^X}{0.5 \left[(1 + \alpha + \alpha^X)(1 + \delta) - \beta^X(1 - \delta) \right]} \right)^{\frac{1}{\gamma}},$$

which is smaller than c_h iff

$$\frac{1 + \alpha + \alpha^X}{0.5 \left[(1 + \alpha + \alpha^X)(1 + \delta) - \beta^X(1 - \delta) \right]} < 1$$

using eq. (A.4) this can be rewritten into:

$$1 + \alpha + \alpha^X < -\beta^X$$

By definition, $\alpha, \alpha^X > 0$, therefore this inequality never holds. This implies that for sufficiently small distances between c_h and c_l (i.e. when c_l converges to c_h) the high standard maximizes aggregate utility for the case (1,1). Overall, assumption 2 assures existence and uniqueness (as shown in section A.3) of the cutoff $\bar{c}_l$.

A.4.3 Cutoff for the low standard: case (0,0)

For the case (0,0), where both groups dissociate from society, the government implements the low standard when $U^{(1,1)}(c_h) < U^{(0,0)}(c_l)$. Setting $U^{(1,1)}(c_h) = U^{(0,0)}(c_l)$ and solving for c_l delivers the cutoff:

$$\tilde{c}_l = \left(\frac{M\Phi c_h^{-\gamma} + A^X}{(1+\alpha)(1+\delta)0.5\Phi} \right)^{-\frac{1}{\gamma}}. \quad (\text{A.8})$$

For $c_l < \tilde{c}_l$ the low standard maximizes aggregate welfare. Obviously, this cutoff is larger than zero and for the range of c_l we are interested in, namely in which identification with society at large is beneficial, $U^{(0,0)}(c_l) < U^{(1,1)}(c_l)$, and hence, $\tilde{c}_l < \bar{c}_l$. As shown in appendix A.3, $U^{(0,0)}(\tilde{c}_l)$ decreases monotonically in c_l , therefore $\tilde{c}_l$ is a unique cutoff that lies between 0 and c_h .

A.4.4 Cutoff ordering

In the following, we derive under which conditions the cutoff ordering $\hat{c}_l < \tilde{c}_l < \bar{c}_l$ applies.²⁴ From appendix A.3 and A.4.3, we already know that aggregate utility decreases in c_l for the parametrization assumed in assumption 2 and that $\tilde{c}_l < \bar{c}_l$. Therefore, it is sufficient to show that $U^{(1,1)}(c_h) < U^{(1,1)}(c_l)$ for $c_l = \hat{c}_l$ to prove that $\hat{c}_l < \tilde{c}_l < \bar{c}_l$. From appendix A.4.2, we already know that this reduces to:

$$(1 + \alpha + \alpha^X)c_h^{-\gamma} < \hat{c}_l^{-\gamma}0.5 \left[(1 + \alpha + \alpha^X)(1 + \delta) - \beta^X(1 - \delta) \right].$$

Plugging in eq. (A.5) and solving for c_h gives:

$$c_h > \left(\frac{(1 + \alpha + \alpha^X)\Phi \left(\beta^X(1 - \delta) - \alpha^X(1 + \delta) \right)}{[(1 + \alpha + \alpha^X)(1 + \delta) - \beta^X(1 - \delta)] A^X} \right)^{\frac{1}{\gamma}}. \quad (\text{A.9})$$

This implies that for all values of c_h that satisfy this condition, the cutoff ordering $\hat{c}_l < \tilde{c}_l < \bar{c}_l$ occurs. The condition is consistent with the lower bound on c_h stated in assumption 3.

²⁴ Note that only for this ordering there exists a range of values for c_l where the free trade agreement becomes welfare-maximizing when social coherence is maintained. In contrast, for the alternative ordering $\bar{c}_l < \tilde{c}_l < \hat{c}_l$ the agreement is welfare-maximizing only for values of c_l where no broad identification occurs.

A.5 Proof of Proposition 3

A.5.1 Fragmentation of caring consumers

We are interested in the case where – apart from identification with the own group – fully congruent caring consumers (fraction $1 - \eta$ with $\tilde{\theta} = 1$) identify with the NGO, while partially congruent caring consumers (fraction η with $\tilde{\theta} = \theta < 1$) identify with society at large. We now derive conditions that allow for this interior solution – i.e. the fragmentation of caring consumers as opposed to the corner solutions where either all caring consumers remain identified only with the NGO or only with society at large – in cost range $c_l^F \in (\tilde{c}_l, \bar{c})$. Caring consumers are indifferent between identification with society at large and identification with the NGO when the last two terms in eq. (7) are equalized:

$$\tilde{\theta}A^N = A^X + \alpha^X \bar{\nu}^X(\hat{c}_l^N) - \beta^X(1 - \lambda_e) \left[\nu_n(\hat{c}_l^N) - \nu_e(\hat{c}_l^N) \right].$$

Plugging in $\bar{\nu}^X(\hat{c}_l^N)$, $\lambda_e = 0.5$ for $i = u, e$, $\nu_n(\hat{c}_l^N)$ and $\nu_e(\hat{c}_l^N)$, this can be rewritten into

$$\tilde{\theta}A^N = A^X - 0.5\Phi(\hat{c}_l^N)^{-\gamma} \left(\beta^X(1 - \delta) - \alpha^X(1 + \delta) \right)$$

and solved for cutoff $\hat{c}_l^N$

$$\hat{c}_l^N = \left(\frac{0.5\Phi \left(\beta^X(1 - \delta) - \alpha^X(1 + \delta) \right)}{A^X - \tilde{\theta}A^N} \right)^{\frac{1}{\gamma}}. \quad (\text{A.10})$$

Caring consumers identify with the NGO (society at large) for $c_l^F < \hat{c}_l^N$ ($c_l^F > \hat{c}_l^N$). Assumption 2 and assumption 4 ensure that the threshold $\hat{c}_l^N$ (eq. (A.10)) is positive. Moreover, assumption 4 implies that $\hat{c}_l^N$ lies above the dissociation cutoff $\hat{c}_l$ (eq. (A.5)), i.e. in the cost range where identification with society is beneficial and caring consumers are willing to identify with society at all. Moreover, from eq. (A.10) follows that $\hat{c}_l^N|_{\tilde{\theta}=\theta} < \hat{c}_l^N|_{\tilde{\theta}=1}$. For $c_l^F \in (\hat{c}_l^N|_{\tilde{\theta}=\theta}, \hat{c}_l^N|_{\tilde{\theta}=1})$, a fragmentation of caring consumers occurs; i.e., caring consumers that are fully congruent to the NGO prototype remain identified with the NGO only and partially congruent caring consumers remain only identified with society. To show that a polarization of caring consumers can generate the TTIP-case, it is sufficient to prove that the cost range $c_l^F \in (\hat{c}_l^N|_{\tilde{\theta}=\theta}, \hat{c}_l^N|_{\tilde{\theta}=1})$ can cover the cost range $c_l^F \in (\tilde{c}_l, \bar{c}_l)$ where the TTIP-case may occur. Hence, we consider the limits of eq. (A.10):

$$\begin{aligned} \lim_{\tilde{\theta} \rightarrow 0} \left(\frac{0.5\Phi \left(\beta^X(1 - \delta) - \alpha^X(1 + \delta) \right)}{A^X - 0 \cdot A^N} \right)^{\frac{1}{\gamma}} &= \left(\frac{0.5\Phi \left(\beta^X(1 - \delta) - \alpha^X(1 + \delta) \right)}{A^X} \right)^{\frac{1}{\gamma}} = \hat{c}_l, \\ \lim_{\tilde{\theta} \rightarrow 1, A^N \rightarrow A^X} \left(\frac{0.5\Phi \left(\beta^X(1 - \delta) - \alpha^X(1 + \delta) \right)}{A^X - 1 \cdot A^X} \right)^{\frac{1}{\gamma}} &= \left(\frac{0.5\Phi \left(\beta^X(1 - \delta) - \alpha^X(1 + \delta) \right)}{0} \right)^{\frac{1}{\gamma}} = \infty. \end{aligned}$$

If $\tilde{\theta} \rightarrow 0$, eq. (A.10) converges to $\hat{c}_l$. From this follows that for a sufficiently low θ the cutoff $\hat{c}_l^N|_{\tilde{\theta}=\theta}$ lies below $\tilde{c}_l$, i.e., the lower bound of range b . If $\tilde{\theta} \rightarrow 1$ and $A^N \rightarrow A^X$, eq. (A.10) converges to ∞ . Hence, for a sufficiently large A^N the cutoff $\hat{c}_l^N|_{\tilde{\theta}=1}$ lies above $\bar{c}_l$ representing the upper bound of range b .

A.5.2 Derivation of the altered ratification cutoff

The altered social identification leads to a shift of the ratification cutoff which is determined by $U^{(1,1,1)}(c_h) = U^{(1,\eta,1-\eta)}(\bar{c}_l')$:

$$\bar{c}_l' = \left[\frac{0.5\Phi \left[(1+\alpha)(1+\delta) - (1+\eta)0.5 \left(\beta^X(1-\delta) - \alpha^X(1+\delta) \right) \right]}{(1+\alpha+\alpha^X)\Phi(c_h)^{-\gamma} + (1-\eta)0.5A^X + 0.5A^N\eta\theta} \right]^{\frac{1}{\gamma}}. \quad (\text{A.11})$$

The cost ranking implies that $c_l^F \in (0, c_h)$. To prove that $\bar{c}_l' \in (0, c_h)$, we consider the limits of c_l^F . We solve eq. (A.11) for c_h and compare $U^{(1,1,1)}(c_h)$ to $U^{(1,\eta,1-\eta)}(c_l^F)$ with $c_l^F = \{0, c_h\}$ for a given c_h to find the optimal standard when c_l^F converges to its limits.

$$\lim_{c_l^F \rightarrow 0} \left(\frac{1+\alpha+\alpha^X}{0.5c_l^{-\gamma} [(1+\alpha)(1+\delta) - (1+\eta)0.5(\beta^X(1-\delta) - \alpha^X(1+\delta))] - (1-\eta)0.5A^X - 0.5A^N\eta\theta} \right)^{\frac{1}{\gamma}} = 0,$$

as the denominator goes to infinity. Since $c_h > 0$, the low standard is preferred when c_l^F approaches zero.

$$\lim_{c_l^F \rightarrow c_h} \left(\frac{1+\alpha+\alpha^X}{0.5c_h^{-\gamma} [(1+\alpha)(1+\delta) - (1+\eta)0.5(\beta^X(1-\delta) - \alpha^X(1+\delta))] - (1-\eta)0.5A^X - 0.5A^N\eta\theta} \right)^{\frac{1}{\gamma}},$$

which is smaller than c_h iff

$$c_h^\gamma > \frac{1+\alpha+\alpha^X}{0.5c_h^{-\gamma} [(1+\alpha)(1+\delta) - (1+\eta)0.5(\beta^X(1-\delta) - \alpha^X(1+\delta))] - (1-\eta)0.5A^X - 0.5A^N\eta\theta}.$$

Since assumption 2 implies that $(1+\alpha)(1+\delta) - (1+\eta)0.5(\beta^X(1-\delta) - \alpha^X(1+\delta)) > 0$, this can be rewritten into

$$-c_h^\gamma 0.5 \left((1-\eta)A^X + A^N\eta\theta \right) > 0.5(1-\delta)(1+\alpha) + [1 - 0.25(1+\eta)(1+\delta)]\alpha^X + 0.25(1+\eta)\beta^X(1-\delta).$$

This inequality never holds as the left-hand side is negative, while the right-hand side is positive. This implies that when c_l^F converges towards c_h , aggregate utility is maximized by the high standard. Hence, for sufficiently small distances between c_h and c_l^F the government prefers to implement the high production standard given that consumers identify with society at large. Overall, assumption 2 ensures existence as well as uniqueness (as shown in section A.3) of the cutoff $\bar{c}_l'$.

A.5.3 Cutoff ordering

We now determine conditions for a cutoff ordering that allows for the TTIP- and CETA-case to appear. First, we know that $U^{(1,\eta,1-\eta)}(c_l^F) = U^{(1,0,1)}(c_l^F)$ at $c_l^F = \hat{c}_l^N|_{\tilde{\theta}=\theta}$ and we show in section A.5.1 that $\hat{c}_l < \hat{c}_l^N|_{\tilde{\theta}=\theta}$; i.e., at $\hat{c}_l^N|_{\tilde{\theta}=\theta}$ benefits from identifying with society exceed dissonance costs. This implies that $U^{(1,\eta,1-\eta)}(c_l^F) = U^{(1,0,1)}(c_l^F) > U^{(0,0,1)}(c_l^F)$ at $\hat{c}_l^N|_{\tilde{\theta}=\theta}$. Considering fig. 3, hence, the curve $U^{(0,0,1)}(c_l^F)$ intersects for a lower level of c_l^F with the horizontal line $U^{(1,1,1)}(c_h)$ than $U^{(1,\eta,1-\eta)}(c_l^F)$ and we can conclude that $\tilde{c}_l < \bar{c}_l'$. Second, the fragmentation of caring consumers and the corresponding identity losses result in a decrease of aggregate welfare, i.e., $U^{(1,\eta,1-\eta)}(c_l^F) < U^{(1,1,1)}(c_l^F)$. Therefore, $\bar{c}_l' < \bar{c}_l$ applies. Knowing that $\tilde{c}_l < \bar{c}_l' < \bar{c}_l$ and along the lines of the argument in section A.4.4, a sufficient condition for the cutoff ordering $\hat{c}_l < \tilde{c}_l < \bar{c}_l' < \bar{c}_l$ is that $U^{(1,1,1)}(c_h) < U^{(1,\eta,1-\eta)}(c_l^F)$ at $c_l^F = \hat{c}_l$. From the derivation of eq. (A.11) we already know that this boils down to:

$$(1 + \alpha + \alpha^X)\Phi c_h^{-\gamma} + 0.5 \left((1 - \eta)A^X + A^N\eta\theta \right) < 0.5\Phi(\hat{c}_l)^{-\gamma} \left[(1 + \alpha + (1 + \eta)0.5\alpha^X)(1 + \delta) - (1 + \eta)0.5\beta^X(1 - \delta) \right] .$$

Plugging in eq. (A.5) gives:

$$(1 + \alpha + \alpha^X)\Phi c_h^{-\gamma} + 0.5 \left((1 - \eta)A^X + A^N\eta\theta \right) < 0.5\Phi \left(\frac{0.5\Phi \left[\beta^X(1 - \delta) - \alpha^X(1 + \delta) \right]}{A^X} \right)^{-1} \left[(1 + \alpha + (1 + \eta)0.5\alpha^X)(1 + \delta) - (1 + \eta)0.5\beta^X(1 - \delta) \right] ,$$

which can be solved for c_h :

$$c_h > \left(\frac{(1 + \alpha + \alpha^X)\Phi \left[\beta^X(1 - \delta) - \alpha^X(1 + \delta) \right]}{A^X \left[(1 + \alpha + \alpha^X)(1 + \delta) - \beta^X(1 - \delta) \right] - 0.5A^N\eta\theta \left[\beta^X(1 - \delta) - \alpha^X(1 + \delta) \right]} \right)^{\frac{1}{\gamma}} . \quad (\text{A.12})$$

This implies that the cutoff ordering $\hat{c}_l < \tilde{c}_l < \bar{c}_l'$ occurs for all values of c_h above the threshold provided in (A.12). Given assumption 2, the numerator in eq. (A.12) is positive. Therefore, assumption 2 implies that there exists an interior threshold for the cost level c_h that allows for the TTIP- and the CETA-case to occur iff:

$$\beta^X < \frac{(1 + \alpha + (1 + 0.5\eta\theta)\alpha^X)(1 + \delta)}{(1 + 0.5\eta\theta)(1 - \delta)} . \quad (\text{A.13})$$

This condition is less strict than the upper bound for β^X stated in assumption 2 and therefore always fulfilled given our parameter restrictions.

We now show that eq. (A.12) is more restrictive than eq. (A.9) and therefore represents the first element of assumption 3. This means that we seek to show that the right-hand side of eq. (A.12) is larger than the right-hand side of eq. (A.9). This boils down to the

following condition:

$$0 > -0.5A^N\eta\theta \left[\beta^X(1-\delta) - \alpha^X(1+\delta) \right] .$$

Since the right-hand side is negative, the inequality above always holds and (A.12) is more restrictive than (A.9). The intuition is the following: $U^{(1,1,1)}(c_h) < U^{(1,\eta,(1-\eta))}(c_l^F) < U^{(1,1,1)}(c_l^F)$ for the range $c_l^F \in (\hat{c}_l, \bar{c}_l')$.²⁵ Therefore, in fig. 3, having the horizontal line $U^{(1,1,1)}(c_h)$ to lie not only below $U^{(1,1,1)}(c_l^F)$ at $c_l^F = \hat{c}_l$, but below the smaller utility level $U^{(1,\eta,(1-\eta))}$ requires a higher c_h – which shifts the horizontal line further downwards. Assumption 3 states the lower bound of c_h given by (A.12) assures the cutoff ordering $\hat{c}_l < \tilde{c}_l < \bar{c}_l' < \bar{c}_l$. Note that Assumption 2 assures the existence of the lower bound.

Moreover, we have to ensure consistency of assumption 3, i.e. that the lower bound is actually smaller than the upper bound for c_h . This implies that:

$$A^N \frac{0.5\eta\theta \left[\beta^X(1-\delta) - \alpha^X(1+\delta) \right]}{(1-\delta)(M-\beta^X)} < A^X .$$

With assumption 4, a sufficient condition for this condition to hold is that the scaling factor of A^N on the left-hand side of the inequality above is smaller than one. This condition delivers the upper bound of assumption 2, which therefore assures that there exists a range for the cost level c_h (stated in assumption 3) that allows for the TTIP- and the CETA-case to occur if:

$$\beta^X < \frac{(1+\alpha+\alpha^X)(1-\delta) + \alpha^X 0.5\eta\theta(1+\delta)}{(1-\delta)(1+0.5\eta\theta)} . \quad (\text{A.14})$$

A.6 Corner solutions (1,1,0) and (1,0,1)

For completeness, we also derive conditions that allow for the TTIP- and CETA-case to occur for the corner solutions (1, 1, 0) and (1, 0, 1).

A.6.1 Case (1,1,0)

The TTIP-case and the CETA-case can arise in the (1, 1, 0) corner solution (all consumers identify with society and none with the NGO) when $U^{(1,1,1)}(c_h) < U^{(1,1,0)}(c_l^F)$ at $c_l^F = \hat{c}_l$:

$$(1+\alpha+\alpha^X)\Phi c_h^{-\gamma} + 0.5A^N(\eta\theta + (1-\eta)) < 0.5\Phi(\hat{c}_l)^{-\gamma} \left[(1+\alpha+\alpha^X)(1+\delta) - \beta^X(1-\delta) \right] .$$

Plugging in eq. (A.5) gives:

$$(1+\alpha+\alpha^X)\Phi c_h^{-\gamma} + 0.5A^N(\eta\theta + (1-\eta)) < 0.5\Phi \left[\left(\frac{0.5\Phi \left[\beta^X(1-\delta) - \alpha^X(1+\delta) \right]}{A^X} \right)^{\frac{1}{\gamma}} \right]^{-\gamma} \cdot \left[(1+\alpha+\alpha^X)(1+\delta) - \beta^X(1-\delta) \right] .$$

²⁵ Taking eq. (9) it is easy to see that $U^{(1,\eta,(1-\eta))}(c_l^F) < U^{(1,1,1)}(c_l^F)$ for $c_l^F \geq \hat{c}_l$. While line one and two are identical in both cases, line three and four are lower in the $(1, \eta, 1-\eta)$ case since both respectively involve utility from identification only for a fraction of caring consumers.

This implies:

$$c_h > \left(\frac{(1 + \alpha + \alpha^X)\Phi [\beta^X(1 - \delta) - \alpha^X(1 + \delta)]}{A^X [(1 + \alpha + \alpha^X)(1 + \delta) - \beta^X(1 - \delta)] - 0.5A^N (\eta\theta + (1 - \eta)) [\beta^X(1 - \delta) - \alpha^X(1 + \delta)]} \right)^{\frac{1}{\gamma}}, \quad (\text{A.15})$$

which is larger than eq. (A.12) and therefore more restrictive on c_h than assumption 3. This implies that c_h has to be larger than for the interior solution in order to allow for the TTIP- and CETA-case.

A.6.2 Case (1,0,1)

The TTIP-case and the CETA-case can arise in the (1, 1, 0) corner solution (all consumers identify with the NGO and none with society) when $U^{(1,1,1)}(c_h) < U^{(1,0,1)}(c_l^F)$ at $c_l^F = \hat{c}_l$:

$$(1 + \alpha + \alpha^X)\Phi c_h^{-\gamma} + A^X < 0.25\Phi(\hat{c}_l)^{-\gamma} [(1 + \alpha + \alpha^X)(1 + \delta) - \beta^X(1 - \delta)]$$

Plugging in eq. (A.5) gives:

$$(1 + \alpha + \alpha^X)\Phi c_h^{-\gamma} + A^X < 0.5\Phi \left[\left(\frac{0.5\Phi [\beta^X(1 - \delta) - \alpha^X(1 + \delta)]}{A^X} \right)^{\frac{1}{\gamma}} \right]^{-\gamma} [(1 + \alpha + \alpha^X)(1 + \delta) - \beta^X(1 - \delta)]. \quad (\text{A.16})$$

This implies

$$c_h > \left(\frac{(1 + \alpha + \alpha^X)\Phi [\beta^X(1 - \delta) - \alpha^X(1 + \delta)]}{A^X [(1 + \alpha + 2\alpha^X)(1 + \delta) - 2\beta^X(1 - \delta)]} \right)^{\frac{1}{\gamma}}, \quad (\text{A.17})$$

which is larger than eq. (A.9) and therefore more restrictive on c_h than assumption 3. This implies that c_h has to be larger than for the interior solution in order to allow for the TTIP- and CETA-case. Moreover, its existence requires a stricter upper bound of β^X than assumption 2, replacing its second element by $\frac{(1+\alpha+2\alpha^X)(1+\delta)}{2(1-\delta)}$.

B Asymmetric population shares

In this section, we rely on a graphical illustration to highlight the main changes in the model when considering unequal population shares. We refer the formal analysis of this case to the Online Appendix, where we characterize ‘old-world’ and ‘new-world’ equilibria with unequal population shares and show that the TTIP- and CETA-case can still arise.

There are two major differences to our analysis with $\lambda_n = \lambda_e$: first, the government puts different weights on the utilities of the two consumer groups. Second, asymmetric population shares affect the consumer decision about identification with society at large. The reason is that dissonance costs increase in the distance between the individual and the group average. When an individual is part of the majority group, it is also closer to the

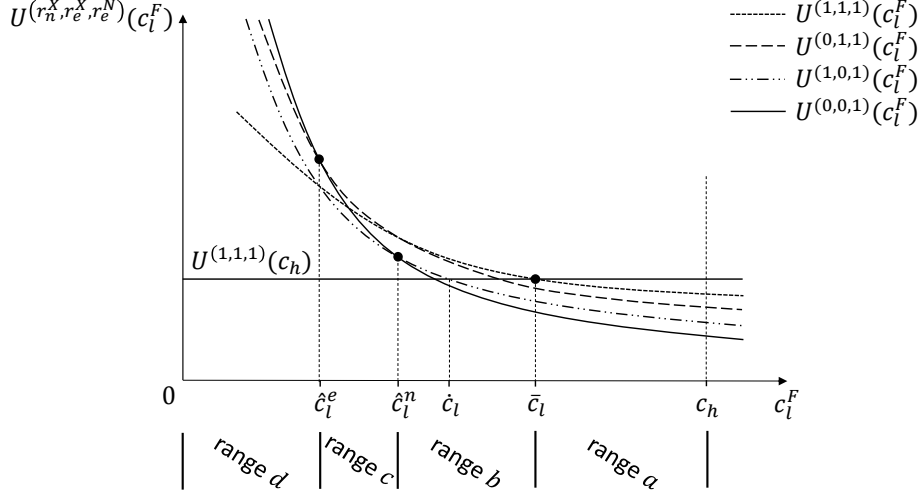


Figure B.1: ‘New world’ equilibrium when $\lambda_n < \lambda_e$

average and is therefore prone to identification, while the minority group will dissociate for relatively low levels of inequality (see equations 6 and 7).

Figure B.1 illustrates the implications of this asymmetry in the identification decision for the case $\lambda_n < \lambda_e$.²⁶ Now, there exist two different dissociation cutoffs; $\hat{c}_l^n$ below which non-caring consumers dissociate from society and $\hat{c}_l^e$ below which caring consumers dissociate. Moreover, mixed cases of (1,0,1) and (0,1,1), where only one group identifies with society at large, may arise. The graph in fig. B.1 therefore features four curves (compared to two curves in our baseline model, see fig. 1). The four curves imply four ratification cutoffs – one for each identification regime – determined by the intersections of the four curves with the horizontal line representing aggregate utility $U^{(1,1,1)}(c_h)$.

To keep the presentation clear, we highlight the two ratification cutoffs that are relevant for the case that we want to analyze, i.e. $\bar{c}_l$ and $\dot{c}_l$, the latter is determined by $U^{(1,0,1)}(c_l^F) = U^{(1,1,1)}(c_h)$. Like for equal population shares, there exists a range $\hat{c}_l^n < c_l^F < \bar{c}_l^n$ (range b) in which the ratification of the trade agreement maximizes aggregate welfare and social cohesion of society is maintained. Within range b, there is a cutoff $\dot{c}_l > \hat{c}_l^n$ above which the trade agreement only maximizes aggregate welfare when the larger group (here: caring consumers) identifies with society at large. Hence, the range $\dot{c}_l < c_l^F < \bar{c}_l$ represents the cost range in which the TTIP-case can occur. If caring consumers split and polarize between the NGO and society at large, aggregate welfare decreases in the initial cases (1, 1, 1) and (0, 1, 1) leading to $U^{(1,\eta,(1-\eta))}(c_l^F)$ and $U^{(0,\eta,(1-\eta))}(c_l^F)$. This in turn induces a downward shift of the ratification cutoff to $\bar{c}_l'$ (see fig. 3). Similar to the case with equal population shares, for the range $c_l^F \in (\bar{c}_l', \bar{c}_l)$, an initially welfare-maximizing trade agreement is rejected (TTIP-case). For the range $c_l^F \in (\hat{c}_l^e, \bar{c}_l')$, a trade agreement is still welfare-maximizing, however, gains are lower than expected and society is polarized (CETA-case). The condition for the TTIP- and CETA-case boils down to $U^{(1,1,1)}(c_h) < U^{(1,\eta,(1-\eta))}(c_l^F)$ at $c_l^F = \hat{c}_l^e$, which holds for a variation of assumption 3 (proof: see the Online Appendix).

²⁶ The graphical illustration of the $\lambda_n > \lambda_e$ case looks identical, except that the curves $U^{(0,1,1)}(c_l^F)$ and $U^{(1,0,1)}(c_l^F)$ are interchanged, and therefore also the cutoffs $\hat{c}_l^n$ and $\hat{c}_l^e$ are as well.